Approximation of improper prior by vague priors

Christele Bioche, Pierre Druilhet

Submitted

APPROXIMATION OF IMPROPER PRIOR BY VAGUE PRIORS

BY CHRISTELE BIOCHE AND PIERRE DRUILHET

Clermont Université

Abstract. We propose a convergence mode for prior distributions which allows a sequence of probability measures to have an improper limiting measure. We define a sequence of vague priors as a sequence of probability measures that converges to a non-informative prior. We consider some cases where vague priors have necessarily large variances and other cases where they have not. We give some constructions of vague priors that approximate the Haar measures or the Jeffreys priors. Then, we study the consequences of the convergence of prior distributions on the posterior analysis. We also revisit the Jeffreys-Lindley paradox.

1. Introduction. The bayesian analysis relies on the Bayes' formula obtained from a prior distribution on the parameter and a conditional probability for the observations. By a formal approach, the Bayes formula may be extended to improper priors, i.e. measures with infinite mass. When no prior information is available, several approaches such as flat priors (Laplace, 1812), Jeffreys' priors (Jeffreys, 1946), reference priors (Berger et al, 2009) or Haar's measures (Eaton, 1989) often lead to improper priors. However, the use of improper prior may cause some problems such as improper posterior priors or undesirable behaviour in hypothesis testing. So, many authors prefer to replace these improper priors by vague priors, i.e. probability measures that aim to represent very few knowledge on the parameter. However, the definition of a vague prior is often "vague". Sometimes, it is defined as a prior with large variance or with pdf having a high spread without reference to any improper prior. Sometimes, it is defined as a prior such that the posterior estimator is close to the estimator obtained from a given improper prior.

Consider for example the standard gaussian model $X|\theta \sim \mathcal{N}(\theta, 1)$. Without prior knowledge, we may choose $\Pi = \lambda_{\mathbb{R}}$ which corresponds to the Harr measure, the Jeffreys prior and also the flat prior. In that case, the Bayes estimator is also the frequentist ML estimator. This improper prior may be

MSC 2010 subject classifications: Primary 62F15; secondary 62E17,62F03,60B10

Keywords and phrases: approximation of improper priors, conjugate priors, convergence of prior, Jeffreys-Lindley paradox, Jeffreys prior, noninformative priors, vague priors

replaced by the vague prior $\Pi_n = \mathcal{N}(0, n)$ with n large. This prior has the two features described above: it has a large variance and the Bayes estimator converges to the estimator obtained from the flat prior when n goes to $+\infty$. It seems then that the distribution $\mathcal{N}(0, n)$ converges to the Lebesgue measure, but there is no usual topology such that this convergence is possible, mainly because the limit of probability measures cannot have a total mass greater than 1. Some other questions arise from this example : does any prior with large variance give the same limiting estimator ? Is the apparent convergence to the Lebesgue measure related to this specific statistical model or is it more intrinsic, i.e. independent from the statistical model ?

The aim of this paper is to give an answer to these questions by proposing a new convergence mode for prior distributions. This convergence mode corresponds to the topology of a quotient space on the Radon measures which arises naturally in the bayesian framework. In Section 2, we define this convergence mode. We show that a sequence of vague priors is related to at most one improper prior, whatever the statistical model is, which means that a sequence of vague priors is associated to a specific prior if any. We also show that there is no discontinuity between proper and improper distributions in the sense that any improper distribution can be approximated by proper distributions and conversely, any proper distribution can be approximated by improper distributions. At last, we propose some construction of vague priors. In Section 3, we give some conditions on the likelihood to derive convergence of posterior distributions and bayesian estimators from the convergence of prior distributions. In Section 4, we give some example of convergence of vague priors. In Section 5, we revisit the Jeffreys-Lindley paradox is the light of our convergence mode.

We warn the reader that the term "vague" is used in two completely independent meanings, even if the aim of the paper is to establish some links between the two notions. The first one is the notion of vague prior discussed above and the second one is the so-called vague topology and related vague convergence on the space of Radon measures (see Definition B.8).

2. Definition, properties and examples of *q*-vague convergence. Let X be a random variable and assume that $X|\theta \sim P_\theta,\ \theta \in \Theta$. We assume that Θ is a locally compact Hausdorff space that is second countable. In practice, Θ is often $\mathbb{R}$, $\mathbb{R}^p, p > 1$, or a countable set. In the bayesian paradigm, a prior distribution Π is given on Θ. This prior may be proper or improper. In this paper, we always assume that Π belongs to the space of non-zero positive Radon measures on Θ, denoted by $\mathcal{R}$. We denote by π the density function w.r.t. some σ-finite measure. If Π is a probability measure, we can use the Bayes Formula to write the posterior density function w.r.t. the same measure μ:

$$\pi(\theta|x) = \frac{f(x|\theta)\ \pi(\theta)}{\int_\Theta f(x|\theta)\pi(\theta)\mathrm{d}\mu(\theta)} \tag{1}$$

where $f(x|\theta)$ is the likelihood function. If Π is an improper measure but $\int_\Theta f(x|\theta)\pi(\theta)\mathrm{d}\mu(\theta) < +\infty$, we can formally applied Formula (1) to get a posterior probability. Now, if we replace Π by $\alpha\Pi$, for $\alpha > 0$, we obtain the same posterior distribution, which means that the prior distribution is defined up to within a scalar factor. So, it is natural to define the equivalence relation $\sim$ on $\mathcal{R}$ by:

$$\Pi \sim \Pi' \iff \exists\ \alpha > 0 \text{ such that } \Pi = \alpha\Pi'. \tag{2}$$

Then, it is natural to define $\overline{\mathcal{R}}$, the quotient space of $\mathcal{R}$ w.r.t. the equivalence relation $\sim$. If we consider prior, or posterior, distributions in the quotient space $\overline{\mathcal{R}}$ rather than in $\mathcal{R}$, then it is natural to write

$$\pi(\theta|x) \propto f(x|\theta)\ \pi(\theta) \tag{3}$$

in place of (1), which is usual in the bayesian litterature. Writing (3), it doesn't matter if the posterior distribution is proper or not.

We denote by $\mathcal{C}_K(\Theta)$ the space of real-valued continuous functions on Θ with compact support and by $\mathcal{C}_K^+(\Theta)$ the positive functions in $\mathcal{C}_K(\Theta)$. When there is no ambiguity on the space, they will be simply denoted by $\mathcal{C}_K$ or $\mathcal{C}_K^+$. We also introduce the notations $\mathcal{C}_b(\Theta)$ for the space of bounded continuous functions on Θ, and $\mathcal{C}_0(\Theta)$ for the space of continuous functions g such that for all $\varepsilon > 0$, there exists a compact $K \subset \Theta$ such that for all $\theta \in K^c$, $g(\theta) < \varepsilon$. We use the notations $\Pi(h) = \int_\Theta h\mathrm{d}\Pi$ where h is a measurable real-valued function, and $|\Pi| = \Pi(1) = \int_\Theta \mathrm{d}\Pi$, the total mass of Π. We denote by $\mathcal{P}$ the space of probability measures on Θ, $\mathcal{M}^b$ the space of non-zero positive finite Radon measures on Θ and $\mathcal{I}$ the space of positive Radon measures on Θ with infinite mass. Note that probability measures are finite measures so $\mathcal{P} \subset \mathcal{M}^b$ and $\mathcal{R} = \mathcal{M}^b \bigcup \mathcal{I}$. Moreover, in the quotient space, $\mathcal{P}$ and $\mathcal{M}^b$ are equivalents.

2.1. *Convergence of prior distribution sequences.* We propose here a convergence mode for sequences of priors which corresponds to the standard quotient topology on $\overline{\mathcal{R}}$ derived from the vague topology on $\mathcal{R}$ (see Definition A.1).

Definition 2.1. *A sequence of non-zero positive Radon measures* $\{\Pi_n\}_{n\in\mathbb{N}}$ *is said to converge q-vaguely to* $\Pi \in \mathcal{R}$, *or to approximate* Π, *if* $\overline{\Pi}_n \xrightarrow[n\to+\infty]{} \overline{\Pi}$.

Now, we give a characterization of q-vague convergence by using vague convergence.

Proposition 2.2. *Let* $\{\Pi_n\}_{n\in\mathbb{N}}$ *and* Π *be in* $\mathcal{R}$. *The sequence* $\{\Pi_n\}_{n\in\mathbb{N}}$ *converges q-vaguely to* Π *iff there exists a sequence of positive real numbers* $\{a_n\}_{n\in\mathbb{N}}$ *such that* $\{a_n\Pi_n\}_{n\in\mathbb{N}}$ *converges vaguely to* Π.

Proof. See appendix. □

When $\Theta = \{\theta_i\}_{i\in I}$ is a discrete set with $I \subset \mathbb{N}$, we give an easy-to-check characterization of the q-vague convergence. The continuous case will be treated in the section 2.2.

Corollary 2.3. *Let* $\{\Pi_n\}_{n\in\mathbb{N}}$ *and* Π *be in* $\mathcal{R}$ *on* $\Theta = \{\theta_i\}_{i\in I}$, $I \subset \mathbb{N}$. *The sequence* $\{\Pi_n\}_{n\in\mathbb{N}}$ *converges q-vaguely to* Π *iff there exists a sequence of positive real numbers* $\{a_n\}_{n\in\mathbb{N}}$ *such that for all* $\theta \in \Theta$, $a_n\Pi_n(\theta) \xrightarrow[n\to+\infty]{} \Pi(\theta)$.

Proof. It is a direct consequence of Proposition 2.2 and Lemma B.10. □

Example 2.4. *Consider* $\Theta = \mathbb{N}$ *and* $\Pi_n = \mathcal{U}(\{0,1,...,n\})$, *then* $\Pi_n(\theta) = \frac{1}{n+1}\,\mathbb{1}_{\{0,1,...,n\}}(\theta)$. *Put* $a_n = n+1$, *then, for* $\theta \in \mathbb{N}$, $a_n\Pi_n(\theta) = \mathbb{1}_{\{0,1,...,n\}}(\theta) \xrightarrow[n\to+\infty]{} 1$. *So,* $\{\mathcal{U}(\{0,1,...,n\})\}_{n\in\mathbb{N}}$ *converges q-vaguely to the counting measure.*

In this article, we define a sequence $\{\Pi_n\}_{n\in\mathbb{N}}$ of vague priors as a sequence of probability measures that converges q-vaguely to an improper measure Π in $\mathcal{R}$. Note that if Θ is a compact space, every positive Radon measure is a finite measure and q-vague convergence is equivalent to vague convergence.

The following proposition shows that a sequence of prior measures cannot converge q-vaguely to more than one limit (up to within a scalar factor).

Theorem 2.5. *Let* $\{\Pi_n\}_{n\in\mathbb{N}}$ *be a sequence in* $\mathcal{R}$ *such that* $\{\Pi_n\}_{n\in\mathbb{N}}$ *converges q-vaguely to both* Π_a *and* Π_b, *then necessarily* $\Pi_a \sim \Pi_b$.

Proof. This is a direct consequence of Proposition A.2 that states that $\overline{\mathcal{R}}$ is a Hausdorff space. However, we give here a direct proof that does not involve abstract topological concept.
Assume that $\{\Pi_n\}_{n\in\mathbb{N}}$ converges q-vaguely to both Π_a and Π_b. From Proposition 2.2, there exist two sequences of positive scalars $\{a_n\}_{n\in\mathbb{N}}$ and $\{b_n\}_{n\in\mathbb{N}}$

such that $\{a_n\Pi_n\}_{n\in\mathbb{N}}$, resp. $\{b_n\Pi_n\}_{n\in\mathbb{N}}$, converges vaguely to Π_a, resp. Π_b. We have to prove that $\Pi_b = \alpha\Pi_a$ for some positive scalar α. Since $\Pi_a \neq 0$ and $\Pi_b \neq 0$, there exist h_a and h_b in $\mathcal{C}_K^+$ such that $\Pi_a(h_a) > 0$ and $\Pi_b(h_b) > 0$. Put $h_0 = h_a + h_b$, we have $\Pi_a(h_0) > 0$ and $\Pi_b(h_0) > 0$. Moreover, $a_n\Pi_n(h_0) \xrightarrow[n\to+\infty]{} \Pi_a(h_0)$ and $b_n\Pi_n(h_0) \xrightarrow[n\to+\infty]{} \Pi_b(h_0)$. So, there exists N such that for $n \geqslant N$, $a_n\Pi_n(h_0) > 0$ and $b_n\Pi_n(h_0) > 0$. For any h in $\mathcal{C}_K$ and $n > N$, $\frac{\Pi_n(h)}{\Pi_n(h_0)} = \frac{a_n\Pi_n(h)}{a_n\Pi_n(h_0)} \xrightarrow[n\to+\infty]{} \frac{\Pi_a(h)}{\Pi_a(h_0)}$ and $\frac{\Pi_n(h)}{\Pi_n(h_0)} = \frac{b_n\Pi_n(h)}{b_n\Pi_n(h_0)} \xrightarrow[n\to+\infty]{} \frac{\Pi_b(h)}{\Pi_b(h_0)}$. By uniqueness of limits in $\mathbb{R}$, $\frac{\Pi_a(h)}{\Pi_a(h_0)} = \frac{\Pi_b(h)}{\Pi_b(h_0)}$. Therefore, from Proposition B.5, $\Pi_a = \frac{\Pi_a(h_0)}{\Pi_b(h_0)}\Pi_b$, i.e. $\Pi_a \sim \Pi_b$. □

LEMMA 2.6. *Let $\{\Pi_n\}_{n\in\mathbb{N}}$ be a sequence of probability measures and $\{a_n\}_{n\in\mathbb{N}}$ a sequence of positive scalars such that $\{a_n\Pi_n\}_{n\in\mathbb{N}}$ converges vaguely to Π in $\mathcal{R}$. If Π is improper, then necessarily $a_n \xrightarrow[n\to+\infty]{} +\infty$.*

PROOF. Let $\{\Pi_n\}_{n\in\mathbb{N}}$ be a sequence of probabilities and Π be an infinite positive Radon measure. We assume that $\{a_n\Pi_n\}_{n\in\mathbb{N}}$ converges vaguely to Π so, from Lemma B.6, we have $\Pi(\Theta) \leqslant \liminf\ a_n\Pi_n(\Theta)$. And for all $n \in \mathbb{N}$, $\Pi_n(\Theta) = 1$ so $\Pi(\Theta) \leqslant \liminf\ a_n$. Moreover, $\Pi(\Theta) = +\infty$ so $\liminf\ a_n = +\infty$. The result follows. □

For a measure Π and a measurable function g, we define the mesure $g\Pi$ by $g\Pi(f) = \Pi(gf) = \int f(\theta)g(\theta)d\Pi(\theta)$ for any f whenever the integrals are defined; $g\Pi$ is also denoted $g\,d\Pi$ or $\Pi \circ g^{-1}$ by some authors.

PROPOSITION 2.7. *Let $\{\Pi_n\}_{n\in\mathbb{N}}$ be a sequence in $\mathcal{R}$ which converges q-vaguely to $\Pi \in \mathcal{R}$. If g is a non-negative continuous function such that $\Pi(g) > 0$, then $\{g\Pi_n\}_{n\in\mathbb{N}}$ converges q-vaguely to $g\Pi$.*

PROOF. Assume that $\{\Pi_n\}_{n\in\mathbb{N}}$ converges q-vaguely to Π. From Proposition 2.2, there exists a sequence of positive scalars $\{a_n\}_{n\in\mathbb{N}}$ such that $\{a_n\Pi_n\}_{n\in\mathbb{N}}$ converges vaguely to Π. So, for any $h \in \mathcal{C}_K$, $a_n\Pi_n(h) \xrightarrow[n\to\infty]{} \Pi(h)$. Since g is a continuous function, $gh \in \mathcal{C}_K$ and $a_n\Pi_n(gh) \xrightarrow[n\to+\infty]{} \Pi(gh)$. But $\Pi_n(gh) = g\Pi_n(h)$ and $\Pi(gh) = g\Pi(h)$. So, $\{a_n g\Pi_n\}_{n\in\mathbb{N}}$ converges vaguely to $g\Pi$, or equivalently $\{g\Pi_n\}_{n\in\mathbb{N}}$ converges q-vaguely to $g\Pi$. □

The following lemma shows that, when a sequence of vague priors is used to approximate an improper prior, the mass tends to concentrate outside any compact set.

PROPOSITION 2.8. *Let $\{\Pi_n\}_{n\in\mathbb{N}}$ be a sequence of probability measures which converges q-vaguely to an improper prior Π. Then, for any compact K in Θ, $\Pi_n(K) \xrightarrow[n\to+\infty]{} 0$, and consequently, $\Pi_n(K^c) \xrightarrow[n\to+\infty]{} 1$.*

PROOF. From Proposition 2.2, there exists $\{a_n\}_{n\in\mathbb{N}}$ such that $a_n\Pi_n(h) \xrightarrow[n\to+\infty]{} \Pi(h)$ for any h in $\mathcal{C}_K$. From Lemma 2.6, $a_n \xrightarrow[n\to+\infty]{} \infty$ whereas $\Pi(h) < +\infty$, so $\Pi_n(h) \xrightarrow[n\to+\infty]{} 0$. Let K_0 be a compact set in Θ. From Lemma B.4, there exists a function $h \in \mathcal{C}_K$ such that $\mathbb{1}_{K_0} \leqslant h$. So $\Pi_n(K_0) \leqslant \Pi_n(h)$ and $\Pi_n(K_0) \xrightarrow[n\to+\infty]{} 0$. Since $\Pi_n(K_0) + \Pi_n(K_0^c) = 1$ for all $n \in \mathbb{N}$, thus $\Pi_n(K_0^c) \xrightarrow[n\to+\infty]{} 1$. □

When Θ is an interval, the following proposition gives the limiting repartition of the mass when the median is a constant.

COROLLARY 2.9. *Let $\{\Pi_n\}_{n\in\mathbb{N}}$ be a sequence of probabilities on $]a,b[$ where $-\infty \leqslant a < b \leqslant +\infty$. We assume that for all n, $\mathrm{med}(\Pi_n) = m \in]a,b[$ and that $\{\Pi_n\}_{n\in\mathbb{N}}$ converges q-vaguely to an improper Radon measure Π. Then, for any $c \in]a,b[$, $\lim_{n\to\infty} \Pi_n(]a,c[) = \frac{1}{2}$ and $\lim_{n\to\infty} \Pi_n(]c,b[) = \frac{1}{2}$.*

PROOF. We only give the proof for $\Pi_n(]a,c[)$. Two cases are considered.

- Assume that $c < m$. For all n, $\Pi_n(]a,c[) + \Pi_n([c,m]) + \Pi_n(]m,b[) = 1$. But, for all n, $\Pi_n(]m,b[) \leqslant \frac{1}{2}$ and, from Proposition 2.8, $\lim_{n\to\infty} \Pi_n([c,m]) = 0$. So $\lim_{n\to\infty} \Pi_n(]a,c[) \geqslant \frac{1}{2}$. Moreover, for all n, $\Pi_n(]a,c[) \leqslant \Pi_n(]a,m[) \leqslant \frac{1}{2}$. So $\lim_{n\to\infty} \Pi_n(]a,c[) = \frac{1}{2}$.
- Assume that $c > m$. For all n, $\Pi_n(]a,c[) = \Pi_n(]a,m[) + \Pi_n([m,c[)$ but $\Pi_n(]a,m[) \leqslant \frac{1}{2}$ and $\lim_{n\to\infty} \Pi_n([m,c[) \leqslant \lim_{n\to\infty} \Pi_n([m,c]) = 0$ from Proposition 2.8. So, for all n, $\lim_{n\to\infty} \Pi_n(]a,c[) \leqslant \frac{1}{2}$. But we also have $\Pi_n(]a,c[) = \Pi_n(]a,m]) + \Pi_n(]m,c[) \geqslant \Pi_n(]a,m]) \geqslant \frac{1}{2}$.

□

Choosing c close to a or b shows that the total mass concentrate equally around a and b. Note that, in Corollary 2.9, we may replace $\mathrm{med}(\Pi_n) = m$ by $\mathrm{med}(\Pi_n) \in [m_1, m_2]$ with $a < m_1 < m_2 < b$.

COROLLARY 2.10. *Under the same notations and assumptions of Corollary 2.9, we have three different cases for the limit of the expectation:*

- *if $-\infty < a$ and $b = +\infty$ then $\mathbb{E}_{\Pi_n}(\theta) \xrightarrow[n\to+\infty]{} +\infty$.*

- *if* $a = -\infty$ *and* $b < +\infty$ *then* $\mathbb{E}_{\Pi_n}(\theta) \xrightarrow[n\to+\infty]{} -\infty$.
- *if* $-\infty < a < b < +\infty$ *then* $\mathbb{E}_{\Pi_n}(\theta) \xrightarrow[n\to+\infty]{} \frac{a+b}{2}$.

PROOF. • Assume that $-\infty < a$ and $b = +\infty$. For b' such that $m < b' < b$, $\mathbb{E}_{\Pi_n}(\theta) = \int_{]a,m[} \theta d\Pi_n(\theta) + \int_{[m,b']} \theta d\Pi_n(\theta) + \int_{]b',b[} \theta d\Pi_n(\theta)$. So $\lim_{n\to\infty} \mathbb{E}_{\Pi_n}(\theta) \geqslant \lim_{n\to\infty} (a\Pi_n(]a,m[) + m\Pi_n([m,b']) + b'\Pi_n(]b',b[))$. By Proposition 2.8, $\lim_{n\to\infty} \Pi_n([m,b']) = 0$. Moreover, by Corollary 2.9, $\lim_{n\to\infty} \Pi_n(]b',b[) = \lim_{n\to\infty} \Pi_n(]a,m[) = \frac{1}{2}$. So, $\lim_{n\to\infty} \mathbb{E}_{\Pi_n}(\theta) \geqslant \frac{1}{2}(a+b')$ for all $b' > m$. The result follows.
The proof is similar for the case $a = -\infty$ and $b < +\infty$.

- Now, assume that $-\infty < a < b < +\infty$.
For $0 < \varepsilon < (b-a)/2$, $\mathbb{E}_{\Pi_n}(\theta) = \int_{]a,a+\varepsilon[} \theta d\Pi_n(\theta) + \int_{[a+\varepsilon,b-\varepsilon]} \theta d\Pi_n(\theta)$ $+ \int_{]b-\varepsilon,b[} \theta d\Pi_n(\theta)$. We have
    - ▷ $a\Pi_n(]a,a+\varepsilon[) \leqslant \int_{]a,a+\varepsilon[} \theta d\Pi_n(\theta) \leqslant (a+\varepsilon)\Pi_n(]a,a+\varepsilon[)$
    - ▷ $(a+\varepsilon)\Pi_n([a+\varepsilon,b-\varepsilon]) \leqslant \int_{[a+\varepsilon,b-\varepsilon]} \theta d\Pi_n(\theta) \leqslant (b-\varepsilon)\Pi_n([a+\varepsilon,b-\varepsilon])$
    - ▷ $(b-\varepsilon)\Pi_n(]b-\varepsilon,b[) \leqslant \int_{]b-\varepsilon,b[} \theta d\Pi_n(\theta) \leqslant b\Pi_n(]b-\varepsilon,b[)$

  Now take the limit when n goes to infinity. From Proposition 2.8 for the second line and from Corollary 2.9 for the first and the third lines, we get after summing, $\frac{1}{2}(a+b-\varepsilon) \leqslant \lim_{n\to\infty} \mathbb{E}_{\Pi_n}(\theta) \leqslant \frac{1}{2}(a+b+\varepsilon)$. Since these inequalities hold for any small ε, $\lim_{n\to\infty} \mathbb{E}_{\Pi_n}(\theta) = \frac{1}{2}(a+b)$.

□

The following theorem states that there is no discontinuity between proper and improper prior in the quotient space $\overline{\mathcal{R}}$.

PROPOSITION 2.11. *Any improper Radon measure may be approximated by a sequence of probability measures and conversely, any proper prior may be approximated by a sequence of improper Radon measures.*

PROOF. We have $\mathcal{R} = \mathcal{M}^b \bigcup \mathcal{I}$ so $\overline{\mathcal{R}} = \overline{\mathcal{M}^b} \bigcup \overline{\mathcal{I}} = \overline{\mathcal{P}} \bigcup \overline{\mathcal{I}}$.

- Let us show that $\mathrm{Adh}(\overline{\mathcal{P}}) = \overline{\mathcal{R}}$.
Consider $\overline{\Pi}$ in $\overline{\mathcal{R}}$ and $\{K_n\}_{n\in\mathbb{N}}$ an increasing sequence of compacts such that $\Theta = \bigcup_{n\in\mathbb{N}} K_n$. Then $\Pi_n = \Pi\mathbb{1}_{K_n}$ is in $\mathcal{M}^b$ so, $\overline{\Pi}_n \in \overline{\mathcal{P}}$. Moreover, $\{\Pi_n\}_{n\in\mathbb{N}}$ converges vaguely, so q-vaguely, to Π.
- We now show that $\mathrm{Adh}(\overline{\mathcal{I}}) = \overline{\mathcal{R}}$.
Let $\overline{\Pi}$ be in $\overline{\mathcal{R}}$. Consider the sequence $\Pi_n = \Pi + \alpha_n\Pi'$ where $\Pi' \in \mathcal{I}$

and $\{\alpha_n\}_{n\in\mathbb{N}}$ is a decreasing sequence which converges to 0. Then, for all $n \in \mathbb{N}$, $\overline{\Pi}_n \in \overline{\mathcal{I}}$ and $\{\Pi_n\}_{n\in\mathbb{N}}$ converges q-vaguely to Π.

□

2.2. *Density functions and q-vague convergence.* In this section, we establish several sufficient conditions for the q-vague convergence of $\{\Pi_n\}_{n\in\mathbb{N}}$ to Π through their pdf. Denote by $\pi_n(\theta) = \frac{\mathrm{d}\Pi_n}{\mathrm{d}\mu}(\theta)$, resp $\pi(\theta) = \frac{\mathrm{d}\Pi}{\mathrm{d}\mu}(\theta)$, the pdf of Π_n, resp Π, w.r.t. some σ-finite measure μ. When Θ is continuous and μ is the Lebesgue measure, then π_n and π are the standard pdf. When Θ is discrete and μ is the counting measure, then $\pi(\theta_0) = \Pi(\theta = \theta_0)$.

THEOREM 2.12. *Let $\{\Pi_n\}_{n\in\mathbb{N}}$ and Π be in $\mathcal{R}$. Assume that:*
1) there exists a sequence of positive real numbers $\{a_n\}_{n\in\mathbb{N}}$ such that the sequence $\{a_n\pi_n\}_{n\in\mathbb{N}}$ converges pointwise to π,
2) $\{a_n\pi_n(\theta)\}_{n\in\mathbb{N}}$ is non-decreasing for all $\theta \in \Theta$.
Then, $\{\Pi_n\}_{n\in\mathbb{N}}$ converges q-vaguely to Π.

PROOF. Let h be in $\mathcal{C}_K^+(\Theta)$. The sequence $\{a_n\, h\, \pi_n\}_{n\in\mathbb{N}}$ is a non-decreasing sequence of non-negative functions. So, by monotone convergence theorem, $\lim\limits_{n\to\infty} \int a_n h(\theta)\pi_n(\theta)\mathrm{d}\mu(\theta) = \int \lim\limits_{n\to\infty} a_n h(\theta)\pi_n(\theta)\mathrm{d}\mu(\theta) = \int h(\theta)\pi(\theta)\mathrm{d}\mu(\theta)$. Thus, from Lemma B.11, $\{\Pi_n\}_{n\in\mathbb{N}}$ converges q-vaguely to Π.

□

EXAMPLE 2.13. *Let $\Theta = \mathbb{R}$, $\Pi_n = \mathcal{N}(0, n^2)$ and $\Pi = \lambda_{\mathbb{R}}$ the Lebesgue measure on $\mathbb{R}$. The corresponding pdf w.r.t. $\mu = \lambda_{\mathbb{R}}$ are $\pi_n(\theta) = \frac{1}{\sqrt{2\pi}n}e^{-\frac{\theta^2}{2n^2}}$ and $\pi(\theta) = 1$. Put $a_n = \sqrt{2\pi}n$, $n \in \mathbb{N}^*$. Then, $\{a_n\pi_n\}_{n\in\mathbb{N}^*}$ is an increasing sequence which converges pointwise to 1. Thus, $\{\mathcal{N}(0, n^2)\}_{n\in\mathbb{N}^*}$ converges q-vaguely to the Lebesgue measure $\lambda_{\mathbb{R}}$.*

EXAMPLE 2.14. *Let $\Theta = \mathbb{R}$, $\Pi_n = \mathcal{U}([-n, n])$, the Uniform distribution on $[-n, n]$, and $\Pi = \lambda_{\mathbb{R}}$ the Lebesgue measure on $\mathbb{R}$. The corresponding pdf w.r.t. $\mu = \lambda_{\mathbb{R}}$ are $\pi_n(\theta) = \frac{1}{2n}\mathbb{1}_{[-n,n]}(\theta)$ and $\pi(\theta) = 1$. Put $a_n = 2n$, $n \in \mathbb{N}^*$. Then, $\{a_n\pi_n\}_{n\in\mathbb{N}^*}$ is an increasing sequence which converges pointwise to 1. So $\{\mathcal{U}([-n, n])\}_{n\in\mathbb{N}^*}$ converges q-vaguely to the Lebesgue measure $\lambda_{\mathbb{R}}$.*

These two exemples will be generalized in Section 2.5.1. Note that, from Theorem 2.5, both $\{\mathcal{N}(0, n^2)\}_{n\in\mathbb{N}}$ and $\{\mathcal{U}([-n, n])\}_{n\in\mathbb{N}}$ cannot converge to another limiting measure than the Lebesgue measure (up to within a scalar factor).

THEOREM 2.15. *Let $\{\Pi_n\}_{n\in\mathbb{N}}$ and Π be in $\mathcal{R}$. Assume that:*
1) there exists a sequence of positive real numbers $\{a_n\}_{n\in\mathbb{N}}$ such that the sequence $\{a_n\pi_n\}_{n\in\mathbb{N}}$ converges pointwise to π,
2) there exists a function $g:\Theta\to\mathbb{R}^+$ such that, for all compact set K, $g\mathbb{1}_K$ is μ-integrable and for all $n\in\mathbb{N}$ and $\theta\in\Theta$, $a_n\pi_n(\theta)<g(\theta)$.
Then, $\{\Pi_n\}_{n\in\mathbb{N}}$ converges q-vaguely to Π.

PROOF. Let h be in $C_K^+(\Theta)$. Then $a_n\pi_n(\theta)h(\theta)\leqslant \parallel h\parallel \mathbb{1}_K g(\theta)$ where $||h|| = \max_{\theta\in\Theta} h(\theta)$. Since $\parallel h\parallel \mathbb{1}_K g(\theta)$ is μ-integrable, by dominated convergence theorem, $\int a_n\pi_n(\theta)h(\theta)\mathrm{d}\mu(\theta) \xrightarrow[n\to+\infty]{} \int \pi(\theta)h(\theta)\mathrm{d}\mu(\theta)$. □

In the following results, we assume that the dominating measure μ is a positive Radon measure, e.g. the Lebesgue measure. In that case, condition 2) in Theorem 2.15 can be replaced by a simpler condition.

COROLLARY 2.16. *Let $\{\Pi_n\}_{n\in\mathbb{N}}$ and Π be in $\mathcal{R}$, and μ is a non-zero positive Radon measure. Assume that:*
1) there exists a sequence of positive real numbers $\{a_n\}_{n\in\mathbb{N}}$ such that the sequence $\{a_n\pi_n\}_{n\in\mathbb{N}}$ converges pointwise to π,
2') there exists a continuous function $g:\Theta\to\mathbb{R}^+$ and $N\in\mathbb{N}$ such that for all $n>N$ and $\theta\in\Theta$, $a_n\pi_n(\theta)<g(\theta)$.
Then, $\{\Pi_n\}_{n\in\mathbb{N}}$ converges q-vaguely to Π.

PROOF. If μ is a non-zero positive Radon measure and g is a continuous function, then $g\mathbb{1}_K$ is μ-integrable. Indeed, $\int_K g\mathrm{d}\mu \leqslant \max_{\theta\in K} g(\theta)\mu(K)$. The result follows from Theorem 2.15. □

EXAMPLE 2.17. *Let $\Theta=\mathbb{R}_+$ and $\Pi_n=\gamma(\alpha_n,\beta_n)$ the Gamma distributions with $(\alpha_n,\beta_n)\xrightarrow[n\to\infty]{}(0,0)$. With $\mu=\lambda_{\mathbb{R}_+}$, $\pi_n(\theta)=\frac{1}{\Gamma(\alpha_n,\beta_n)}\theta^{\alpha_n-1}$ $e^{-\beta_n\theta}$. Put $a_n=\Gamma(\alpha_n,\beta_n)$. Then $a_n\pi_n(\theta)=\theta^{\alpha_n-1}e^{-\beta_n\theta}$ and $\{a_n\pi_n(\theta)\}_{n\in\mathbb{N}}$ converges to $\frac{1}{\theta}$. Put $g(\theta)=\frac{1}{\theta}\mathbb{1}_{[0,1]}(\theta)+\mathbb{1}_{]1,+\infty[}(\theta)$. The sequence $\{\alpha_n\}_{n\in\mathbb{N}}$ goes to 0 so there exists N such that for all $n>N$, $\alpha_n<1$. So, for $n>N$ and for $\theta>0$, $a_n\pi_n(\theta)\leqslant\theta^{\alpha_n-1}\leqslant g(\theta)$. Since g is a continuous function on $\mathbb{R}_+^*$, from Corollary 2.16, $\{\Pi_n\}_{n\in\mathbb{N}}$ converges q-vaguely to $\frac{1}{\theta}\,\mathrm{d}\theta$.*

The following result will be useful to establish a result in Section 2.5.3.

THEOREM 2.18. *Let $\{\Pi_n\}_{n\in\mathbb{N}}$ and Π be in $\mathcal{R}$, and μ is a non-zero positive Radon measure. Assume that:*
1) there exists a sequence of positive real numbers $\{a_n\}_{n\in\mathbb{N}}$ such that the

sequence $\{a_n\pi_n\}_{n\in\mathbb{N}}$ *converges pointwise to* π,
2") for any compact set K*, there exists a scalar* M *and some* $N \in \mathbb{N}$ *such that for* $n > N$, $\sup_{\theta\in K} a_n\, \pi_n(\theta) < M$.
Then, $\{\Pi_n\}_{n\in\mathbb{N}}$ *converges q-vaguely to* Π.

PROOF. The proof is similar to the proof of Theorem 2.15 with $a_n\pi_n(\theta)h(\theta) \leqslant M\, \sup_{\theta\in K} |h(\theta)|$. □

2.3. *Divergence of variances.* Many authors consider that few knowledge on the parameter is represented by priors with large variance. In this section, we establish some links between q-vague convergence and variances of the prior sequence.

PROPOSITION 2.19. *Let* $\{\Pi_n\}_{n\in\mathbb{N}}$ *be a sequence of probabilities on* $\mathbb{R}$ *such that* $\mathbb{E}_{\Pi_n}(\theta)$ *is a constant. If* $\{\Pi_n\}_{n\in\mathbb{N}}$ *converges q-vaguely to an improper prior* Π*, then* $\lim\limits_{n\to\infty} \mathrm{Var}_{\Pi_n}(\theta) = +\infty$.

PROOF. Since $\mathbb{E}_{\Pi_n}(\theta)$ is constant, $\lim\limits_{n\to\infty} \mathrm{Var}_{\Pi_n}(\theta) = +\infty$ iff $\lim\limits_{n\to\infty} \mathbb{E}_{\Pi_n}(\theta^2) = +\infty$. For any $r > 0$, $\mathbb{E}_{\Pi_n}(\theta^2) \geqslant \int_{[-r,r]^c} \theta^2 \mathrm{d}\Pi_n(\theta)$ so $\mathbb{E}_{\Pi_n}(\theta^2) \geqslant r^2\Pi_n([-r,r]^c)$. From Proposition 2.8, $\lim\limits_{n\to\infty} \Pi_n([-r,r]^c) = 1$ and then $\lim\limits_{n\to\infty} \mathbb{E}_{\Pi_n}(\theta^2) \geq r^2$. Since this holds for any $r > 0$, $\lim\limits_{n\to\infty} \mathbb{E}_{\Pi_n}(\theta^2) = +\infty$. □

EXAMPLE 2.20. *Let* $\{\Pi_n\}_{n\in\mathbb{N}}$ *be a sequence of probabilities with constant mean which approximate the Lebesgue measure* $\lambda_{\mathbb{R}}$*. Then, necessarily,* $\mathrm{Var}_{\Pi_n}(\theta) \xrightarrow[n\to+\infty]{} +\infty$*. This is the case, for example, for* $\{\mathcal{N}(0,n^2)\}_{n\in\mathbb{N}}$ *and* $\{\mathcal{U}([-n,n])\}_{n\in\mathbb{N}}$.

In the following proposition, we consider a sequence of probabilities on an interval of $\mathbb{R}$ with constant median.

PROPOSITION 2.21. *Let* $\{\Pi_n\}_{n\in\mathbb{N}}$ *be a sequence of probabilities on* $]a,+\infty[$, $]-\infty,a[$ *or* $\mathbb{R}$*. Assume that* $\mathrm{med}(\Pi_n)$ *is a constant and that* $\{\Pi_n\}_{n\in\mathbb{N}}$ *converges q-vaguely to an improper prior* Π*. Then,* $\lim\limits_{n\to\infty} \mathrm{Var}_{\Pi_n}(\theta) = +\infty$.

PROOF. Denote $m = \mathrm{med}(\Pi_n)$ and $\mu_n = \mathbb{E}_{\Pi_n}(\theta)$. Then, $\mathrm{Var}_{\Pi_n}(\theta) = \mathbb{E}_{\Pi_n}\left((\theta-\mu_n)^2\right)$.

- Consider the case $\Theta =]a,+\infty[$. From Corollary 2.10, $\mu_n \xrightarrow[n\to\infty]{} +\infty$. So, there exists $N \in \mathbb{N}$ such that for all $n > N$, $\mu_n > m$. Thus, for $n > N$, $\mathbb{E}_{\Pi_n}\left((\theta-\mu_n)^2\right) = \int_{]a,m[} (\theta-\mu_n)^2 \mathrm{d}\Pi_n(\theta) + \int_{[m,+\infty[} (\theta-\mu_n)^2 \mathrm{d}\Pi_n(\theta) \geqslant$

$\int_{]a,m[}(\theta-\mu_n)^2 \mathrm{d}\Pi_n(\theta) \geqslant \frac{1}{2}(\mu_n-m)^2$. But $(\mu_n-m)^2 \xrightarrow[n\to+\infty]{} +\infty$, so $\mathrm{Var}_{\Pi_n}(\theta) \xrightarrow[n\to+\infty]{} +\infty$.

- Consider now the case $\Theta=\mathbb{R}$. For any $c>|m|$, if $\mu_n<m$, $\mathrm{Var}_{\Pi_n}(\theta) \geqslant \int_c^{+\infty}(\theta-\mu_n)^2 \mathrm{d}\Pi_n(\theta) \geqslant \int_c^{+\infty}(\theta-m)^2 \mathrm{d}\Pi_n(\theta) \geqslant \int_c^{+\infty}(c-m)^2 \mathrm{d}\Pi_n(\theta) = (c-m)^2 \Pi_n(]c,+\infty[)$. And, for any $c>|m|$, if $\mu_n>m$, $\mathrm{Var}_{\Pi_n}(\theta) \geqslant \int_{-\infty}^{-c}(c+m)^2 \mathrm{d}\Pi_n(\theta) = (c+m)^2\Pi_n(]-\infty,c[)$. Thus, in all cases, $\mathrm{Var}_{\Pi_n}(\theta) \geqslant \max\left\{(c+m)^2\Pi_n(]-\infty,-c[),(c-m)^2\Pi_n(]c,+\infty[)\right\}$. From Corollary 2.9, $\lim_n \Pi_n(]-\infty,-c[) = \lim_n \Pi_n(]c,+\infty[) = \frac{1}{2}$. So, $\lim_n \mathrm{Var}_{\Pi_n}(\theta) \geqslant \frac{1}{2}\max\left\{(c+m)^2,(c-m)^2\right\}$. Since this inequality holds when c goes to $+\infty$, then $\mathrm{Var}_{\Pi_n}(\theta) \xrightarrow[n\to+\infty]{} +\infty$.

□

We now give a generalization of Proposition 2.21 in which we assume neither the expectation nor the median to be constant.

Proposition 2.22. *Let $\{\Pi_n\}_{n\in\mathbb{N}}$ be a sequence of probabilities on $\theta\in]a,b[$, $-\infty \leqslant a < b \leqslant +\infty$. If there exists c with $a<c<b$ such that $\lim_{n\to+\infty}\Pi_n(]a,c[) = \alpha$ for some $0<\alpha<1$. Then,*

1. $\mathrm{Var}_{\Pi_n}(\theta) \xrightarrow[n\to+\infty]{} +\infty$ *if $a=-\infty$ or $b=+\infty$ or both.*
2. $\mathrm{Var}_{\Pi_n}(\theta) \xrightarrow[n\to+\infty]{} \alpha(1-\alpha)(b-a)^2$ *if $-\infty<a<b<+\infty$.*

Proof. From Proposition 2.8, $\lim_{n\to+\infty}\Pi_n(]a,c[)=\alpha$ for some $c\in]a,b[$ is equivalent to $\lim_{n\to+\infty}\Pi_n(]a,a'[)=\alpha$ for any $a'\in]a,b[$ which is also equivalent to $\lim_{n\to+\infty}\Pi_n(]b',b[)=1-\alpha$ for any $b'\in]a,b[$.

Step 1: For all $n\in\mathbb{N}$, $\mathrm{Var}_{\Pi_n}(\theta)=\frac{1}{2}\int\int(x-y)^2\mathrm{d}\Pi_n(x)\mathrm{d}\Pi_n(y)$. So, for any $a<a'<b'<b$, $\mathrm{Var}_{\Pi_n}(\theta) \geqslant \int\int_{]a,a'[\times]b',b[}(x-y)^2\mathrm{d}\Pi_n(x)\mathrm{d}\Pi_n(y) \geqslant (b'-a')^2\int\int_{]a,a'[\times]b',b[}\mathrm{d}\Pi_n(x)\mathrm{d}\Pi_n(y) = (b'-a')^2\Pi_n(]a,a'[)\Pi_n(]b',b[)$. So $\lim_{n\to+\infty}\mathrm{Var}_{\Pi_n}(\theta) \geqslant (b'-a')^2\alpha(1-\alpha)$ for all a', b' such that $a<a'<b'<b$. Taking $a'\to a$ and $b'\to b$, we get $\lim_{n\to+\infty}\mathrm{Var}_{\Pi_n}(\theta) \geqslant (b-a)^2\alpha(1-\alpha)$ if $-\infty<a<b<+\infty$ and $\lim_{n\to+\infty}\mathrm{Var}_{\Pi_n}(\theta)=+\infty$ if $a=-\infty$ or $b=+\infty$.

Step 2: For any $a<a'<b'<b$, we denote by $A_1=]a,a'[$, $A_2=[a',b']$, $A_3=]b',b[$ and $B_{ij}=A_i\times A_j$, $(i,j)\in\{1,2,3\}^2$. For all $n\in\mathbb{N}$, $\mathrm{Var}_{\Pi_n}(\theta) = \sum_{i,j}\int\int_{B_{ij}}(x-y)^2\mathrm{d}\Pi_n(x)\mathrm{d}\Pi_n(y)$. We have the following inequalities:

- ▷ $\int\int_{B_{11}}(x-y)^2 \mathrm{d}\Pi_n(x)\mathrm{d}\Pi_n(y) \leqslant (a-a')^2\Pi_n(B_{11})$
- ▷ $\int\int_{B_{22}}(x-y)^2 \mathrm{d}\Pi_n(x)\mathrm{d}\Pi_n(y) \leqslant (b'-a')^2\Pi_n(B_{22})$
- ▷ $\int\int_{B_{33}}(x-y)^2 \mathrm{d}\Pi_n(x)\mathrm{d}\Pi_n(y) \leqslant (b-b')^2\Pi_n(B_{11})$
- ▷ $\int\int_{B_{12}\bigcup B_{21}}(x-y)^2 \mathrm{d}\Pi_n(x)\mathrm{d}\Pi_n(y) \leqslant 2(b'-a)^2\Pi_n(B_{12})$
- ▷ $\int\int_{B_{32}\bigcup B_{23}}(x-y)^2 \mathrm{d}\Pi_n(x)\mathrm{d}\Pi_n(y) \leqslant 2(b-a')^2\Pi_n(B_{23})$
- ▷ $\int\int_{B_{31}\bigcup B_{13}}(x-y)^2 \mathrm{d}\Pi_n(x)\mathrm{d}\Pi_n(y) \leqslant 2(b-a)^2\Pi_n(B_{23})$

And, $\Pi_n(B_{11}) = \Pi_n(A_1) \times \Pi_n(A_1) \xrightarrow[n\to+\infty]{} \alpha^2$, $\Pi_n(B_{22}) \xrightarrow[n\to+\infty]{} 0$, $\Pi_n(B_{33}) \xrightarrow[n\to+\infty]{} (1-\alpha)^2$, $\Pi_n(B_{12}) \xrightarrow[n\to+\infty]{} 0$, $\Pi_n(B_{23}) \xrightarrow[n\to+\infty]{} 0$ and $\Pi_n(B_{13}) \xrightarrow[n\to+\infty]{} \alpha(1-\alpha)$.

So, $\lim_{n\to+\infty} \mathrm{Var}_{\Pi_n}(\theta) \leqslant \alpha^2(a-a')^2 + (b-b')^2(1-\alpha)^2 + (b-a)^2\alpha(1-\alpha)$.

When $a' \to a$ and $b' \to b$, we have $\lim_{n\to+\infty} \mathrm{Var}_{\Pi_n}(\theta) \leqslant \alpha(1-\alpha)(b-a)^2$.

Combining with Step 1, we get $\lim_{n\to+\infty} \mathrm{Var}_{\Pi_n}(\theta) = \alpha(1-\alpha)(b-a)^2$ if $-\infty < a < b < +\infty$.

□

See Section 4 for examples of more complex situations.

2.4. *Reparameterization.* In many statistical models, there are several parameterizations of interest. In this section we study the impact of the change of parameterization on the q-vague convergence of prior distributions. Consider a new parameterization $\eta = h(\theta)$ where h is a homeomorphism. We denote by $\widetilde{\Pi}_n = \Pi_n \circ h^{-1}$ and $\widetilde{\Pi} = \Pi \circ h^{-1}$ the prior distribution on η derived from the prior distribution on θ. The following proposition establishes a link between q-vague convergence of $\{\Pi_n\}_{n\in\mathbb{N}}$ and $\{\widetilde{\Pi}_n\}_{n\in\mathbb{N}}$.

PROPOSITION 2.23. *Let $\{\Pi_n\}_{n\in\mathbb{N}}$ be a sequence of priors in $\mathcal{R}$ which converges q-vaguely to Π. Let h be a homeomorphism and consider the parameterization $\eta = h(\theta)$. Then $\{\widetilde{\Pi}_n\}_{n\in\mathbb{N}}$ converges q-vaguely to $\widetilde{\Pi}$.*

PROOF. From the change of variables formula, $\int g(h(\theta))\mathrm{d}\Pi_n(\theta) = \int g(\eta)\, \mathrm{d}\widetilde{\Pi}_n(\eta)$ and $\int g(h(\theta))\mathrm{d}\Pi(\theta) = \int g(\eta)\, \mathrm{d}\widetilde{\Pi}(\eta)$. Morevover, if $\{\Pi_n\}_{n\in\mathbb{N}}$ converges q-vaguely to Π, from Proposition 2.2 there exists $\{a_n\}_{n\in\mathbb{N}}$ such that $\{a_n\Pi_n\}_{n\in\mathbb{N}}$ converges vaguely to Π. Note that for all $g \in \mathcal{C}_K$, $g\circ h \in \mathcal{C}_K$. So, for all $g \in \mathcal{C}_K$, $a_n \int g(h(\theta))\mathrm{d}\Pi_n(\theta) \xrightarrow[n\to\infty]{} \int g(h(\theta))\mathrm{d}\Pi(\theta)$, i.e. $a_n \int g(\eta)\mathrm{d}\widetilde{\Pi}_n(\eta) \xrightarrow[n\to+\infty]{} \int g(\eta)\, \mathrm{d}\widetilde{\Pi}(\eta)$. Thus $\{\widetilde{\Pi}_n\}_{n\in\mathbb{N}}$ converges q-vaguely to $\widetilde{\Pi}$. □

REMARK 2.24. *Proposition 2.23 holds if h is just continuous instead of being a homeomorphism. However, in that case, the q-vague convergence of $\{\widetilde{\Pi}_n\}_{n\in\mathbb{N}}$ does not imply the convergence of $\{\Pi_n\}_{n\in\mathbb{N}}$.*

EXAMPLE 2.25. *Assume that $X|\theta \sim \mathcal{E}(\theta)$ an Exponential distribution and put the prior $\gamma(\frac{1}{n},\frac{1}{n})$ on $\theta \in \mathbb{R}_+^*$. Consider the reparameterization $\eta = h(\theta) = \frac{1}{\theta} = \mathbb{E}(X|\theta)$. We denote $\Pi_n = \gamma(\frac{1}{n},\frac{1}{n})$ and π_n the pdf. For $\theta \sim \gamma(\frac{1}{n},\frac{1}{n})$, $\eta \sim \mathcal{IG}(\frac{1}{n},\frac{1}{n})$, an Inverse Gaussian distribution. We have $\pi_n(\theta) = \theta^{\frac{1}{n}-1}\, e^{-\frac{\theta}{n}}\, \left(\frac{1}{n}\right)^{\frac{1}{n}} \frac{1}{\Gamma(\frac{1}{n},\frac{1}{n})}$. Put $a_n = \Gamma(\frac{1}{n},\frac{1}{n})\frac{1}{(\frac{1}{n})^{\frac{1}{n}}}$, then $a_n\pi_n(\theta) = \theta^{\frac{1}{n}-1}e^{-\frac{\theta}{n}}$. Thus $\{a_n\pi_n(\theta)\}_{n\in\mathbb{N}^*}$ is an increasing sequence and $a_n\pi_n(\theta) \xrightarrow[n\to+\infty]{} \frac{1}{\theta} = \pi(\theta)$ so from Theorem 2.12 $\{\gamma(\frac{1}{n},\frac{1}{n})\}_{n\in\mathbb{N}^*}$ converges q-vaguely to $\frac{1}{\theta}\,\mathrm{d}\theta$. The function h is a homeomorphism so, from Proposition 2.23, $\{\mathcal{IG}(\frac{1}{n},\frac{1}{n})\}_{n\in\mathbb{N}^*}$ converges q-vaguely to $\frac{1}{\eta}\,\mathrm{d}\eta$.*

2.5. *Some constructions of sequences of vague priors.* In this section, we give some examples of construction of sequences of vague priors. In the first two examples, we build sequences which converge q-vaguely to Haar's measures on the groups $(\mathbb{R},+)$ and $(\mathbb{R}_+^*,\times)$. In the third example, we prove the q-vague convergence of a sequence of conjugate priors for exponential families to Jeffreys prior.

2.5.1. *Location models.* The parameter θ is said to be a location parameter if the conditional distribution of $X-\theta$ given θ is the same for all θ. For instance, it's the case when $X|\theta \sim \mathcal{N}(\theta,\sigma^2)$ with σ^2 known. The concerned group is $(\mathbb{R},+)$ where Haar's measure $\lambda_{\mathbb{R}}$ is improper, so we want to approximate her.

PROPOSITION 2.26. *Let Π be a probability measure on $\mathbb{R}$ such that the density π of Π w.r.t. Lebesgue measure $\lambda_{\mathbb{R}}$ is continuous at 0, bounded and satisfy $\pi(0) > 0$. Denote by ϕ_n the application*

$$\begin{array}{cccc} \phi_n & : & \mathbb{R} & \rightarrow & \mathbb{R} \\ & & a & \mapsto & na \end{array}.$$

Then, $\{\Pi \circ \phi_n^{-1}\}_{n\in\mathbb{N}^*}$ converges q-vaguely to $\lambda_{\mathbb{R}}$.

PROOF. Put $\Pi_n = \Pi\circ\phi_n^{-1}$. The density of Π_n w.r.t. Lebesgue measure is $\pi_n(\theta) = |\frac{\partial\phi_n^{-1}}{\partial\theta}|\pi(\phi_n^{-1}(\theta))$. Here, $\pi_n(\theta) = \frac{1}{n}\pi(\frac{\theta}{n})$. Put $a_n = n$, then $a_n\pi_n(\theta) = \pi(\frac{\theta}{n}) \xrightarrow[n\to\infty]{} \pi(0) > 0$ since π is continuous at 0. Moreover, π is bounded so there exists $M > 0$ such that for all $\theta \in \mathbb{R}$ and for all $n \in \mathbb{N}^*$, $\pi(\frac{\theta}{n}) \leqslant$

$M = g(\theta)$. The function g is continuous so, from Corollary 2.16, the result follows. $\square$

We can note that Hartigan (1996) used a dual approach. He reduced the impact of the prior by making the conditional variance σ^2 go to 0, and he gets similarly conditions. He assumes that Π is locally uniform at 0, but it is equivalent to assume that Π is continuous and positive at 0. And we replace his condition "Π tail-bounded" by the condition "Π bounded".

COROLLARY 2.27. *Let Π be a probability measure on $\mathbb{R}$ such that the density π of Π w.r.t. Lebesgue measure $\lambda_{\mathbb{R}}$ is continuous and such that $\pi(0) > 0$. Denote by ϕ_n the application*

$$\begin{array}{cccc} \phi_n & : & \mathbb{R} & \rightarrow & \mathbb{R} \\ & & a & \mapsto & na \end{array}.$$

Then, $\{\Pi \circ \phi_n^{-1}\}_{n\in\mathbb{N}^*}$ converges q-vaguely to $\lambda_{\mathbb{R}}$.

REMARK 2.28. *In Proposition 2.26, we may replace the assumption "π bounded" by "π dominated by a continuous function".*

2.5.2. *Scale models.* The strictly positive parameter θ is said to be a scale parameter if the conditional distribution of $\frac{1}{\theta}X$ given θ is the same for all θ. For instance, it's the case when $X|\theta \sim \mathcal{E}(\theta)$. Here, the concerned group is $(\mathbb{R}_+^*, \times)$ where Haar's measure $\frac{1}{\theta}\lambda_{\mathbb{R}_+^*}$ is improper. The following proposition is the equivalent of the previous one for Haar's measure on $(\mathbb{R}_+^*, \times)$.

PROPOSITION 2.29. *Let Π be a probability measure on $\mathbb{R}$ such that the density π of Π w.r.t. Lebesgue measure $\lambda_{\mathbb{R}_+^*}$ is continuous at 1, bounded and satisfy $\pi(1) > 0$. Denote by ϕ_n the application*

$$\begin{array}{cccc} \phi_n & : & \mathbb{R} & \rightarrow & \mathbb{R} \\ & & a & \mapsto & a^n \end{array}.$$

Then, $\{\Pi \circ \phi_n^{-1}\}_{n\in\mathbb{N}^*}$ converges q-vaguely to $\frac{1}{\theta}\lambda_{\mathbb{R}_+^*}$.

PROOF. Put $\Pi_n = \Pi \circ \phi_n^{-1}$. The density of Π_n w.r.t. Lebesgue measure is $\pi_n(\theta) = |\frac{\partial \phi_n^{-1}}{\partial \theta}|\pi(\phi_n^{-1}(\theta))$. Here, $\pi_n(\theta) = \frac{1}{\theta\, n}\theta^{\frac{1}{n}}\pi(\theta^{\frac{1}{n}})$. Put $a_n = n$, then $a_n\pi_n(\theta) = \theta^{\frac{1}{n}-1}\pi(\theta^{\frac{1}{n}}) \xrightarrow[n\to\infty]{} \frac{1}{\theta}\pi(1)$ since π is continuous at 1. Moreover, π is bounded so there exists $M > 0$ such that for all $\theta \in \mathbb{R}_+^*$, $\pi(\theta) \leqslant M$. Put $g(\theta) = \frac{M}{\theta}\mathbb{1}_{\theta<1} + M\mathbb{1}_{\theta\geqslant 1}$. Then, for all $n \in \mathbb{N}^*$ and $\theta > 0$, $\theta^{\frac{1}{n}-1}\pi(\theta^{\frac{1}{n}}) \leqslant g(\theta)$. The function g is continuous so, from Corollary 2.16, the result follows. $\square$

2.5.3. *Jeffreys conjugate priors (JCPs).* The Jeffreys prior is one of the most popular prior when no information is available, but, in many cases, is improper. Consider that the distribution $X|\theta$ belongs to an exponential family, i.e. $f(x|\theta) = \exp\{\theta \cdot t(x) - \phi(\theta)\}\, h(x)$, for some functions $t(x)$, $h(x)$ and $\phi(\theta)$, and $\theta \in \Theta$, where Θ is an open set in $\mathbb{R}^p$, $p \geq 1$, such that $f(x|\theta)$ is a well defined pdf. We assume that $\phi(\theta)$ and $I_\theta(\theta)$ are continuous. These conditions are satisfied if $t(X)$ is not concentrated on an hyperplane a.s. (see Barndorff-Nielsen, 1978). Druilhet and Pommeret (2012) proposed a class of conjugate priors that aims to approximate the Jeffreys prior and that is invariant w.r.t. smooth reparameterization. The notion of approximation was defined only from an intuitive point of view. We can now give a more rigorous approach by using the q-vague convergence.

Denote by $\pi^J(\theta) = |I_\theta(\theta)|^{1/2}$ the pdf of the Jeffreys prior w.r.t. the Lebesgue measure, where θ is the natural parameter of the exponential family and $I_\theta(\theta)$ is the determinant of Fisher information matrix. The JCPs are defined through their pdf w.r.t. the Lebesgue measure by

$$\pi^J_{\alpha,\beta}(\theta) \propto \exp\{\alpha.\theta - \beta\phi(\theta)\}|I_\theta(\theta)|^{\frac{1}{2}},$$

and for a smooth reparameterization $\theta \to \eta$ by

$$\pi^J_{\alpha,\beta}(\eta) \propto \exp\{\alpha.\theta(\eta) - \beta\phi(\theta(\eta))\}|I_\eta(\eta)|^{\frac{1}{2}}.$$

PROPOSITION 2.30. *Let $\{(\alpha_n, \beta_n)\}_{n\in\mathbb{N}}$ be a sequence of real numbers that converges to $(0,0)$. Then, for the natural parameter θ or for any smooth reparameterization η, $\{\Pi^J_{\alpha_n,\beta_n}\}_{n\in\mathbb{N}}$ converges q-vaguely to Π^J.*

PROOF. Choose a_n such that $a_n\pi^J_{\alpha_n,\beta_n}(\theta) = \exp\{\alpha_n\,\theta - \beta_n\phi(\theta)\}|I_\theta(\theta)|^{\frac{1}{2}}$, which converges pointwise to $|I_\theta(\theta)|^{\frac{1}{2}}$. Put $\gamma_n = (\alpha_n, \beta_n)$ and $\psi(\theta) = (\theta, -\phi(\theta))$. We have $\gamma_n \cdot \psi(\theta) = \alpha_n\,\theta - \beta_n\phi(\theta)$. By Cauchy-Schwarz inequality, $\gamma_n \cdot \psi(\theta) \leq \|\gamma_n\|\ \|\psi(\theta)\|$. Since γ_n converges to $(0,0)$, there exists N such that, for $n > N$, $\|\gamma_n\| < 1$. Let K be a compact set in Θ. By continuity of $\psi(\theta)$, since $\phi(\theta)$ is continuous, and by continuity of $I_\theta(\theta)$, there exist M_1 and M_2 such that, for all $\theta \in K$, $\|\psi(\theta)\| < M_1$ and $|I_\theta(\theta)|^{\frac{1}{2}} < M_2$. Therefore, $a_n\pi^J_{\alpha_n,\beta_n}(\theta) \leq M_2\exp\{M_1\}$. The result follows from Theorem 2.18. □

Even if we have the convergence to the Jeffreys prior, we have no guaranty that $\Pi^J_{\alpha_n,\beta_n}$ is a proper prior and there is no general result to characterize this property such as in Diaconis and Ylvisaker (1979) for usual conjugate priors. In the case of quadratic exponential families (see Morris, 1983), JCPs and usual conjugate priors coincide and correspond to

Normal, Gamma or Beta distribution which are examined throughout this paper. Druilhet and Pommeret (2012) considered inverse gaussian models, which are not quadratic exponential families, are considered. In that case, $f(x;\mu,\lambda) = \left(\frac{\lambda}{2\pi x^3}\right)^{\frac{1}{2}} \exp\left(\frac{-\lambda(x-\mu)^2}{2\mu^2 x}\right) \mathbb{1}_{\{x>0\}}$ where $\mu > 0$ denotes the mean parameter and $\lambda > 0$ stands for the shape parameter. Considering the parameterization $\left(\psi = \frac{1}{\mu}, \lambda\right)$, the JCPs are given by $\pi^J_{\alpha,\beta}(\psi,\lambda) \propto e^{-\frac{\lambda}{2}(\alpha_1\psi^2 - 2\beta\psi + \alpha_2)}$ $\psi^{-\frac{1}{2}}\lambda^{\frac{(\beta-1)}{2}}$. They showed that $\pi^J_{\alpha,\beta}(\psi,\lambda)$ is proper iff $\alpha_1 > 0$, $\alpha_2 > 0$ and $-\frac{1}{2} \leqslant \beta < \sqrt{\alpha_1\alpha_2}$. So, consider the sequences $\alpha_{1,n} = \alpha_{2,n} = \frac{1}{n}$ and $\beta_n = \frac{1}{2n}$. By Proposition 2.30, $\Pi^J_{\alpha_n,\beta_n}(\psi,\lambda)$ is a sequence of proper priors that converge q-vaguely to the Jeffreys prior Π^J.

REMARK 2.31. *For any continuous function g on Θ, we can define $\pi^g_{\alpha,\beta}(\theta) \propto \exp\{\alpha.\theta - \beta\phi(\theta)\}g(\theta)$ and $\pi^g(\theta) = g(\theta)$. Similarly to Proposition 2.30, it can be shown that $\{\Pi^g_{\alpha_n,\beta_n}\}_{n\in\mathbb{N}}$ converges q-vaguely to Π^g.*

3. Convergence of posterior distributions and estimators. Consider the model $X|\theta \sim P_\theta$, $\theta \in \Theta$. We denote by $f(x|\theta)$ the likelihood. The priors $\Pi_n(\theta)$ on Θ represent our prior knowledge. In this section, we study the consequences of the q-vague convergence of $\{\Pi_n\}_{n\in\mathbb{N}}$ on the posterior analysis. We always assume that $\Pi(\cdot|x) \in \mathcal{R}$, which is the case, for example, if $\theta \longmapsto f(x|\theta)$ is continuous and $\int f(x|\theta)\mathrm{d}\Pi(\theta) > 0$.

3.1. *Convergence of posterior distributions.* In this part, we study the q-vague convergence of the sequence of posteriors $\{\Pi_n(\cdot|x)\}_{n\in\mathbb{N}}$ to $\Pi(\cdot|x)$ when $\{\Pi_n\}_{n\in\mathbb{N}}$ converges q-vaguely to Π.

LEMMA 3.1. *Let $\{\Pi_n\}_{n\in\mathbb{N}}$ be a sequence of priors on $\mathbb{R}$ which converges q-vaguely to $\Pi \in \mathcal{R}$. Assume that $\theta \longmapsto f(x|\theta)$ is a non-zero continuous function on Θ. Then $\{\Pi_n(\cdot|x)\}_{n\in\mathbb{N}}$ converges q-vaguely to $\Pi(\cdot|x)$.*

PROOF. This is a direct consequence of Proposition 2.7. $\square$

When the limiting measure $\Pi(\cdot|x)$ is a probability measure, we can establish results about the narrow convergence of $\{\Pi_n(\cdot|x)\}_{n\in\mathbb{N}}$ to $\Pi(\cdot|x)$ instead of q-vague convergence. Before, we give a necessary and sufficient condition for the narrow convergence of a sequence of probabilities which converges q-vaguely to a probability.

PROPOSITION 3.2. *Let $\{\Pi_n\}_{n\in\mathbb{N}}$ and Π be probability measures such that $\{\Pi_n\}_{n\in\mathbb{N}}$ converges q-vaguely to Π. Then $\{\Pi_n\}_{n\in\mathbb{N}}$ converges narrowly to Π iff $\{\Pi_n\}_{n\in\mathbb{N}}$ is tight.*

Proof. Direct part: $\{\Pi_n\}_{n\in\mathbb{N}}$ converges narrowly to Π a probability measure so $\{\Pi_n\}_{n\in\mathbb{N}}$ is tight.
Converse part: Let us show that any subsequence of $\{\Pi_n\}_{n\in\mathbb{N}}$ which converges narrowly converges to Π. From Theorem B.16 there exists a subsequence $\{\Pi_{n_k}\}_{k\in\mathbb{N}}$ of $\{\Pi_n\}_{n\in\mathbb{N}}$ which converges narrowly to some probability measure, say $\widetilde{\Pi}$. Since $\{\Pi_{n_k}\}_{k\in\mathbb{N}}$ is a sequence of probabilities which converges narrowly to $\widetilde{\Pi}$, from Proposition 2.2, $\{\Pi_{n_k}\}_{k\in\mathbb{N}}$ converges q-vaguely to $\widetilde{\Pi}$. So, from Theorem 2.5, $\Pi \sim \widetilde{\Pi}$, but Π and $\widetilde{\Pi}$ are probabilities. So $\Pi = \widetilde{\Pi}$. The result follows from Lemma B.17. □

Theorem 3.3. *Let $\{\Pi_n\}_{n\in\mathbb{N}}$ and Π be in $\mathcal{R}$, $\pi_n(\theta) = \frac{\mathrm{d}\Pi_n}{\mathrm{d}\mu}$ and $\pi(\theta) = \frac{\mathrm{d}\Pi}{\mathrm{d}\mu}$ where μ is a σ-finite measure. Assume that:*
1) there exists a sequence of positive real numbers $\{a_n\}_{n\in\mathbb{N}}$ such that the sequence $\{a_n\pi_n\}_{n\in\mathbb{N}}$ converges pointwise to π,
2) $\{a_n\pi_n(\theta)\}_{n\in\mathbb{N}}$ is non-decreasing for all $\theta \in \Theta$,
3) $\theta \longmapsto f(x|\theta)$ is continuous and positive,
4) $\Pi(\cdot|x)$ is proper.
Then, $\{\Pi_n(\cdot|x)\}_{n\in\mathbb{N}}$ converges narrowly to $\Pi(\cdot|x)$.

Proof. The sequence $\{a_n\, f\, \pi_n\}_{n\in\mathbb{N}}$ is a non-decreasing sequence of non-negative functions. By monotone convergence theorem, $\lim_{n\to\infty} \int a_n f(x|\theta)\, \pi_n(\theta)\, \mathrm{d}\mu(\theta) = \int \lim_{n\to\infty} a_n f(x|\theta)\pi_n(\theta)\mathrm{d}\mu(\theta) = \int f(x|\theta)\pi(\theta)\mathrm{d}\mu(\theta)$. So, $\{a_n\Pi_n(f)\}_{n\in\mathbb{N}}$ converges to $\Pi(f) > 0$. So there exists N such that for all $n > N$, $a_n\Pi_n(f) \geqslant \frac{1}{2}\Pi(f)$. Consider $\{K_m\}_{m\in\mathbb{N}}$ an increasing sequence of compact sets such that $\bigcup K_m = \Theta$. The sequence $\{K_m^c\}_{m\in\mathbb{N}}$ decreases to $\varnothing$ so $\lim_{m\to\infty} \Pi(f\mathbb{1}_{K_m^c}) = 0$. Thus, for all $\varepsilon > 0$, there exists M such that, for all $m \geqslant M$, $\Pi(f\mathbb{1}_{K_m^c}) \leqslant \varepsilon$. So, for all $n > N$, $\frac{f\,\Pi_n(K_M^c)}{\Pi_n(f)} = \frac{f\,a_n\Pi_n(K_M^c)}{a_n\Pi_n(f)} \leqslant \frac{2\,a_n\Pi_n(f\mathbb{1}_{K_M^c})}{\Pi(f)} \leqslant \frac{2\Pi(f\mathbb{1}_{K_M^c})}{\Pi(f)} \leqslant \frac{2\varepsilon}{\Pi(f)}$. The second inequality comes from assumption 3). Thus, $\{\frac{f\,\Pi_n}{\Pi_n(f)}\}_{n\in\mathbb{N}}$ is tight. The result follows from Proposition 3.2. □

Example 3.4. *Assume that $X|\theta \sim \mathcal{N}(\theta, \sigma^2)$, σ^2 known, and put the prior $\Pi_n = \mathcal{N}(0, n^2)$ on θ. Then, $\Pi_n(\theta|x) = \mathcal{N}(\frac{n^2 x}{\sigma^2+n^2}, \frac{\sigma^2 n^2}{\sigma^2+n^2})$. From Example 2.13, the first and the second hypothesis are satisfied and $\{\mathcal{N}(0, n^2)\}_{n\in\mathbb{N}}$ converges q-vaguely to Lebesgue measure $\lambda_{\mathbb{R}}$ so here, $\Pi = \lambda_{\mathbb{R}}$. Moreover, $\theta \longmapsto f(x|\theta)$ is continuous and positive on Θ and $\Pi(\cdot|x) = \mathcal{N}(x, \sigma^2)$ is proper. So, from Theorem 3.3, $\{\mathcal{N}(\frac{n^2 x}{\sigma^2+n^2}, \frac{\sigma^2 n^2}{\sigma^2+n^2})\}_{n\in\mathbb{N}}$ converges narrowly to $\mathcal{N}(x, \sigma^2)$.*

THEOREM 3.5. *Let* $\{\Pi_n\}_{n\in\mathbb{N}}$ *and* Π *be in* $\mathcal{R}$, $\pi_n(\theta) = \frac{\mathrm{d}\Pi_n}{\mathrm{d}\mu}$ *and* $\pi(\theta) = \frac{\mathrm{d}\Pi}{\mathrm{d}\mu}$ *where* μ *is a* σ*-finite measure. Assume that:*
1) there exists a sequence of positive real numbers $\{a_n\}_{n\in\mathbb{N}}$ *such that the sequence* $\{a_n\pi_n\}_{n\in\mathbb{N}}$ *converges pointwise to* π,
2) there exists a continuous function $g : \Theta \to \mathbb{R}^+$ *such that* fg *is* μ*-integrable and for all* $n \in \mathbb{N}$ *and* $\theta \in \Theta$, $a_n\pi_n(\theta) < g(\theta)$,
3) $\theta \longmapsto f(x|\theta)$ *is continuous and positive,*
4) $\Pi(\cdot|x)$ *is proper.*
Then, $\{\Pi_n(\cdot|x)\}_{n\in\mathbb{N}}$ *converges narrowly to* $\Pi(\cdot|x)$.

PROOF. For all n, $a_n f(x|\theta)\pi_n(\theta) \leqslant f(x|\theta)g(\theta)$. Since fg is μ-integrable, by dominated convergence theorem, $\lim\limits_{n\to\infty} \int a_n\, f(x|\theta)\, \pi_n(\theta)\, \mathrm{d}\mu(\theta) = \int \lim\limits_{n\to\infty} a_n$ $f(x|\theta)\, \pi_n(\theta)\, \mathrm{d}\mu(\theta) = \int f(x|\theta)\pi(\theta)\mathrm{d}\mu(\theta)$. Thus, $\{a_n\Pi_n(f)\}_{n\in\mathbb{N}}$ converges to $\Pi(f) > 0$ so there exists N such that for all $n > N$, $a_n\Pi_n(f) \geqslant \frac{1}{2}\Pi(f)$. Consider $\{K_m\}_{m\in\mathbb{N}}$ an increasing sequence of compact sets such that $\bigcup K_m = \Theta$. The sequence $\{K_m^c\}_{m\in\mathbb{N}}$ decreases to $\varnothing$ so $\lim\limits_{m\to\infty} \mu(fg\mathbb{1}_{K_m^c}) = 0$. Thus, for all $\varepsilon > 0$, there exists M such that for all $m \geqslant M$, $\mu(fg\mathbb{1}_{K_m^c}) \leqslant \varepsilon$. So, for all $n > N$, $\frac{f\, a_n\Pi_n(K_M^c)}{a_n\Pi_n(f)} \leqslant \frac{2\, a_n\Pi_n(f\mathbb{1}_{K_M^c})}{\Pi(f)} \leqslant \frac{2\mu(fg\mathbb{1}_{K_M^c})}{\Pi(f)} \leqslant \frac{2\varepsilon}{\Pi(f)}$. Thus, $\{\frac{f\,\Pi_n}{\Pi_n(f)}\}_{n\in\mathbb{N}}$ is tight. The result follows from Proposition 3.2. □

The following result will be useful to explain the Jeffreys-Lindley paradox (see Section 5).

THEOREM 3.6. *Consider a sequence of probabilities* $\{\Pi_n\}_{n\in\mathbb{N}}$ *which converges vaguely to the proper measure* Π. *Assume that:*
1) $\theta \longmapsto f(x|\theta)$ *is continuous and non-negative,*
2) $f(x|\cdot) \in \mathcal{C}_0(\Theta)$.
Then, $\{\Pi_n(\cdot|x)\}_{n\in\mathbb{N}}$ *converges narrowly to* $\Pi(\cdot|x)$.

PROOF. Since Π is a proper measure and $f(\cdot|\theta)$ is a pdf, $\Pi(\cdot|x)$ is a probability. We assume that $\{\Pi_n\}_{n\in\mathbb{N}}$ converges vaguely, and so q-vaguely, to Π and that f satisfies 1). So, from Proposition 2.7, $\{\Pi_n(\cdot|x)\}_{n\in\mathbb{N}}$ converges q-vaguely to $\Pi(\cdot|x)$. From Lemma B.9, $\{\Pi_n(f)\}_{n\in\mathbb{N}}$ converges to $\Pi(f)$. So, there exists N such that for $n > N$, $\Pi_n(f) > \frac{\Pi(f)}{2}$. Moreover, from assumption 2), for all $\varepsilon > 0$, there exists a compact K such that for all $\theta \in K^c$, $f(\theta|x) \leqslant \varepsilon$. Thus, for all $n > N$, $\frac{f\Pi_n(K^c)}{\Pi_n(f)} \leqslant \frac{2\Pi_n(f\mathbb{1}_{K^c})}{\Pi(f)} \leqslant \frac{2\varepsilon}{\Pi(f)}$. Thus, $\{\frac{f\,\Pi_n}{\Pi_n(f)}\}_{n\in\mathbb{N}}$ is tight. The result follows from Proposition 3.2. □

3.2. *Convergence of estimators.* We consider the estimator $\mathbb{E}_{\Pi}(\theta|x)$ of θ which minimizes the quadratic risk. Let us give a proposition about convergence of the sequence of posterior estimators when the sequence of priors $\{\Pi_n\}_{n\in\mathbb{N}}$ converges q-vaguely to Π.

PROPOSITION 3.7. *Let $\{\Pi_n\}_{n\in\mathbb{N}}$ be a sequence of $\mathcal{R}$ which converges q-vaguely to $\Pi \in \mathcal{R}$. Assume that $\theta \longmapsto f(x|\theta)$ is a non-zero continuous function on Θ, and that the family $\{\Pi_n(\cdot|x)\}_{n\in\mathbb{N}}$ is a family of probabilities uniformly integrable (see Definition B.18). Then, $\mathbb{E}_{\Pi_n}(\theta|x) \xrightarrow[n\to+\infty]{} \mathbb{E}_{\Pi}(\theta|x)$.*

PROOF. From Lemma 3.1, $\{\Pi_n(\theta|x)\}_{n\in\mathbb{N}}$ converges q-vaguely to $\Pi(\theta|x)$. For all n, $\Pi_n(\cdot|x)$ and $\Pi(\cdot|x)$ are probability measures and $\{\Pi_n(\cdot|x)\}_{n\in\mathbb{N}}$ uniformly integrable implies that $\{\Pi_n(\cdot|x)\}_{n\in\mathbb{N}}$ is tight. So, from Proposition 3.2, $\{\Pi_n(\theta|x)\}_{n\in\mathbb{N}}$ converges narrowly to $\Pi(\theta|x)$. The result follows from Billingsley (1968, th 5.4). $\square$

We can give an other version of Proposition 3.7 with a condition more restrictive than uniform integrability but easier to check.

THEOREM 3.8. *Let $\{\Pi_n\}_{n\in\mathbb{N}}$ be a sequence of $\mathcal{R}$ which converges q-vaguely to $\Pi \in \mathcal{R}$. Assume that $\theta \longmapsto f(x|\theta)$ is a non-zero continuous function on Θ, and that $\{\Pi_n(\cdot|x)\}_{n\in\mathbb{N}}$ is a family of probabilities such that $\{\mathrm{Var}_{\Pi_n}(\theta|x)\}_n$ is bounded above. Then $\mathbb{E}_{\Pi_n}(\theta|x) \xrightarrow[n\to+\infty]{} \mathbb{E}_{\Pi}(\theta|x)$.*

PROOF. This is a consequence of Proposition B.19 and Proposition 3.7. $\square$

EXAMPLE 3.9. *To continue 3.4, $\mathrm{Var}_{\Pi_n}(\theta|x) = \frac{\sigma^2 n^2}{\sigma^2+n^2}$ is bounded above by σ^2 and the other hypothesis have already been verified in Example 3.4. So, from Proposition 3.7, $\mathbb{E}_{\Pi_n}(\theta|x) \xrightarrow[n\to\infty]{} \mathbb{E}_{\Pi}(\theta)$. Indeed, $\mathbb{E}_{\Pi_n}(\theta) = \frac{n^2 x}{\sigma^2+n^2} \xrightarrow[n\to\infty]{} x = \mathbb{E}_{\Pi}(\theta)$.*

4. Some examples.

4.1. *Poisson distribution.* Here is an exemple where a family of vague priors does not converge. Consider the sequence of Poisson distribution $\Pi_n = \mathcal{P}(n)$, then $\pi_n(\theta) = \exp(-n)\frac{n^\theta}{\theta!}$ w.r.t. the counting measure. Assume that there exists $\Pi \in \mathcal{R}$ such that $\{\Pi_n\}_{n\in\mathbb{N}}$ converges q-vaguely to Π. Then, from Corollary 2.3, there exists a sequence $\{a_n\}_{n\in\mathbb{N}}$ such that for all $\theta \in \Theta$, $a_n\Pi_n(\theta) \xrightarrow[n\to\infty]{} \Pi(\theta)$. Consider $\theta_0 \in \Theta$ such that $\Pi(\theta_0) > 0$. There exists N such that, for all $n > N$, $\Pi_n(\theta_0) > 0$. Consider $\theta > \theta_0$, for all $n > N$, $\frac{\Pi_n(\theta)}{\Pi_n(\theta_0)} = \frac{\theta_0!}{\theta!}n^{\theta-\theta_0}$ and $\frac{\Pi_n(\theta)}{\Pi_n(\theta_0)} \xrightarrow[n\to\infty]{} \frac{\Pi(\theta)}{\Pi(\theta_0)} < +\infty$. But $\frac{\theta_0!}{\theta!}n^{\theta-\theta_0} \xrightarrow[n\to\infty]{} +\infty$. So, there is no $\Pi \in \mathcal{R}$ such that $\{\Pi_n\}_{n\in\mathbb{N}}$ converges q-vaguely to Π.

4.2. *Normal distribution.* We have seen in Example 2.13 that the sequence $\{\mathcal{N}(0,n^2)\}_{n\in\mathbb{N}^*}$ converges q-vaguely to Lebesgue measure $\lambda_{\mathbb{R}}$. Similarly, it can be shown that the limiting measure is the same for $\{\mathcal{N}(\mu_n,n^2)\}_n$ where $\{\mu_n\}_{n\in\mathbb{N}^*}$ is a constant or a bounded sequence. So, we consider now the case where $\mu_n \xrightarrow[n\to\infty]{} +\infty$ by taking $\mu_n = n$.

PROPOSITION 4.1. *We have three cases for the convergence of* $\mathcal{N}(n,\sigma_n^2)$*:*

1. *If* $\frac{n}{\sigma_n^2} \xrightarrow[n\to+\infty]{} +\infty$*, then* $\{\mathcal{N}(n,\sigma_n^2)\}_{n\in\mathbb{N}}$ *doesn't converge q-vaguely.*
2. *If* $\frac{n}{\sigma_n^2} \xrightarrow[n\to+\infty]{} c$ *with* $0 < c < \infty$*, then* $\{\mathcal{N}(n,\sigma_n^2)\}_{n\in\mathbb{N}}$ *converges q-vaguely to* $\exp(c\theta)\mathrm{d}\theta$*.*
3. *If* $\frac{n}{\sigma_n^2} \xrightarrow[n\to+\infty]{} 0$*, then* $\{\mathcal{N}(n,\sigma_n^2)\}_{n\in\mathbb{N}}$ *converges q-vaguely to* $\lambda_{\mathbb{R}}$*.*

PROOF. For all $n > 0$, we denote by $\Pi_n = \mathcal{N}(n,\sigma_n^2)$, and by π_n the pdf w.r.t. the Lebesgue meausure, $\pi_n(\theta) = \frac{1}{\sqrt{2\pi}\sigma_n}\exp(-\frac{(\theta-n)^2}{2\sigma_n^2})$.

1. Put $\widetilde{\pi}_n(\theta) = \exp(-\frac{\theta^2}{2\sigma_n^2} + \frac{\theta\, n}{\sigma_n^2})$. For all n, Π_n and $\widetilde{\Pi}_n$ are equivalents so $\{\Pi_n\}_{n\in\mathbb{N}}$ converges q-vaguely iff $\{\widetilde{\Pi}_n\}_{n\in\mathbb{N}}$ converges q-vaguely. Assume that there exists $\widetilde{\Pi}$ such that $\{\widetilde{\Pi}_n\}_{n\in\mathbb{N}}$ converges q-vaguely to $\widetilde{\Pi}$. Then, there exists a sequence $\{a_n\}_{n\in\mathbb{N}}$ such that $\{a_n\widetilde{\Pi}_n\}_{n\in\mathbb{N}}$ converges vaguely to $\widetilde{\Pi}$. Since $\widetilde{\Pi} \neq 0$, there exists an interval $[A_1,A_2]$ such that $-\infty < A_1 < A_2 < +\infty$ and $0 < \widetilde{\Pi}([A_1,A_2]) < +\infty$. Consider $[B_1,B_2]$ such that $A_2 < B_1 < B_2 < +\infty$. There exists N such that for $n > N$, $\theta \longmapsto -\frac{\theta^2}{2n} + \frac{\theta\, n}{\sigma_n^2}$ is non-decreasing. For a such n, $\widetilde{\Pi}_n([B_1,B_2]) \geqslant (B_2 - B_1)\ \exp(-\frac{B_1}{2\sigma_n^2} + \frac{B_1\, n}{\sigma_n^2})$ and $\widetilde{\Pi}_n([A_1,A_2]) \leqslant (A_2 - A_1)\exp(-\frac{A_2}{2\sigma_n^2} + \frac{A_2\, n}{\sigma_n^2})$. So $\frac{\widetilde{\Pi}_n([B_1,B_2])}{\widetilde{\Pi}_n([A_1,A_2])} \geqslant \frac{B_2-B_1}{A_2-A_1}\exp(C(n))$ with $C(n) = \frac{n(B_1-A_2)}{\sigma_n^2} - \frac{(B_1^2-A_2^2)}{2\sigma_n^2} \geqslant \frac{n(B_1-A_2)}{2\sigma_n^2}$. Thus, $\frac{\widetilde{\Pi}_n([B_1,B_2])}{\widetilde{\Pi}_n([A_1,A_2])} \xrightarrow[n\to\infty]{} +\infty$ but $\frac{\widetilde{\Pi}_n([B_1,B_2])}{\widetilde{\Pi}_n([A_1,A_2])} \xrightarrow[n\to\infty]{} \frac{\widetilde{\Pi}([B_1,B_2])}{\widetilde{\Pi}([A_1,A_2])} < +\infty$. So, $\{\Pi_n\}_{n\in\mathbb{N}}$ doesn't converge q-vaguely.
2. Put $a_n = \frac{1}{\sqrt{2\pi}\sigma_n}\exp(-\frac{n^2}{\sigma_n^2})$. Then $a_n\pi_n(\theta) = \exp(-\frac{\theta^2}{2\sigma_n^2} + \frac{\theta n}{\sigma_n^2}) \xrightarrow[n\to\infty]{} e^{c\theta}$. Moreover, because $\frac{n}{\sigma_n^2} \xrightarrow[n\to+\infty]{} c$, there exists N such that for all $n > N$, $\frac{n}{\sigma_n^2} \in [c-\varepsilon, c+\varepsilon]$. So, for all $n > N$, $\exp(-\frac{\theta^2}{2\sigma_n^2} + \frac{\theta n}{\sigma_n^2}) \leqslant \exp((c+\varepsilon)\theta)$ which is continuous. The result follows from Corollary 2.16.
3. This is the same reasoning as Point 2. with $a_n\pi_n(\theta) \xrightarrow[n\to+\infty]{} 1$ and $a_n\pi_n(\theta) \leqslant 1+\varepsilon$ for all $n > N$ for N large enough.

□

4.3. *Gamma distribution.*

4.3.1. *Approximation of* $\Pi = \frac{1}{\theta}\mathbb{1}_{\theta>0}\mathrm{d}\theta$. We have shown in Example 2.17 that $\{\gamma(\alpha_n,\beta_n)\}_{n\in\mathbb{N}}$ converges q-vaguely to $\frac{1}{\theta}\mathbb{1}_{\theta>0}\mathrm{d}\theta$ if $\{\alpha_n\}_{n\in\mathbb{N}}$ and $\{\beta_n\}_{n\in\mathbb{N}}$ are decreasing sequences which go to 0. Recall that for $\theta\sim\gamma(a,b)$, $\mathbb{E}(\theta)=\frac{a}{b}$ and $\mathrm{Var}(\theta)=\frac{a}{b^2}$. We can see below that the same convergence may be obtained with different convergence of the mean and variance.

- For $\Pi_n=\gamma(\frac{1}{n},\frac{1}{n})$, $\mathbb{E}_{\Pi_n}(\theta)=1$ for all n and $\mathrm{Var}_{\Pi_n}(\theta)=n\to+\infty$.
- For $\Pi_n=\gamma(\frac{1}{n},\frac{1}{\sqrt{n}})$, $\mathbb{E}_{\Pi_n}(\theta)=\frac{1}{\sqrt{n}}\to 0$ and $\mathrm{Var}_{\Pi_n}(\theta)=1$ for all n.
- For $\Pi_n=\gamma(\frac{1}{n},\frac{1}{n^{\frac{1}{3}}})$, $\mathbb{E}_{\Pi_n}(\theta)=n^{-\frac{2}{3}}\to 0$ and $\mathrm{Var}_{\Pi_n}(\theta)=n^{-\frac{1}{3}}\to 0$.
- For $\Pi_n=\gamma(\frac{1}{n},\frac{1}{n^2})$, $\mathbb{E}_{\Pi_n}(\theta)=n\to+\infty$ and $\mathrm{Var}_{\Pi_n}(\theta)=n^3\to+\infty$.
- For $\Pi_n=\gamma(\frac{1}{n},\frac{1}{n^{\frac{2}{3}}})$, $\mathbb{E}_{\Pi_n}(\theta)=n^{-\frac{1}{2}}\to 0$ and $\mathrm{Var}_{\Pi_n}(\theta)=n^{\frac{1}{3}}\to+\infty$.

More generally, if $\liminf\limits_{n\to\infty}\mathbb{E}_{\Pi_n}(\theta)>0$ then $\mathrm{Var}_{\Pi_n}(\theta)\xrightarrow[n\to+\infty]{}+\infty$, since $\mathrm{Var}_{\Pi_n}(\theta)=\frac{\mathbb{E}_{\Pi_n}(\theta)}{\beta_n}$ with $\lim\beta_n=0$.

4.3.2. *Approximation of* $\Pi=\frac{1}{\theta}e^{-\theta}\mathbb{1}_{\theta>0}\mathrm{d}\theta$. Let us show that $\{\gamma(\alpha_n,1)\}_{n\in\mathbb{N}}$ converges q-vaguely to $\frac{1}{\theta}e^{-\theta}\mathbb{1}_{\theta>0}\mathrm{d}\theta$ when $\{\alpha_n\}_{n\in\mathbb{N}}$ goes to 0. Put $\Pi_n=\{\gamma(\alpha_n,1)\}_{n\in\mathbb{N}}$. Then $\pi_n(\theta)=\frac{1}{B(\alpha_n,1)}\ \theta^{\alpha_n-1}e^{-\theta}\mathbb{1}_{\theta>0}$ is the pdf of Π_n. Put $a_n=B(\alpha_n,1)$, then $a_n\pi_n(\theta)=\theta^{\alpha_n-1}e^{-\theta}\mathbb{1}_{\theta>0}$ converges to $\pi(\theta)=\frac{1}{\theta}e^{-\theta}\mathbb{1}_{\theta>0}$. Moreover, because $\{\alpha_n\}_{n\in\mathbb{N}}$ goes to 0, there exists N such that for $n>N$, $\alpha_n<1$. Put $g(\theta)=\frac{1}{\theta}\mathbb{1}_{]0,1]}(\theta)+\mathbb{1}_{]1,+\infty[}(\theta)$. So, for $n>N$ and $\theta>0$, $a_n\pi_n(\theta)\leqslant\theta^{\alpha_n-1}\leqslant g(\theta)$. The function g is continuous so from Corollary 2.16, $\{\gamma(\alpha_n,1)\}_{n\in\mathbb{N}}$ converges q-vaguely to $\frac{1}{\theta}e^{-\theta}\mathbb{1}_{\theta>0}\mathrm{d}\theta$. Since $\alpha_n\xrightarrow[n\to+\infty]{}0$, we necessarily have $\mathbb{E}_{\Pi_n}(\theta)\xrightarrow[n\to+\infty]{}0$ and $\mathrm{Var}_{\Pi_n}(\theta)\xrightarrow[n\to+\infty]{}0$.

4.4. *Beta distribution.* We now treat a more complex example which often appears in literature, see e.g. Tuyl et al (2009). Let X represents the number of successes in N Bernoulli trials, and θ be the probability of a success in a single trial. Thus, $X\in\{0,1,...,N\}$, and $X|\theta\sim\mathcal{B}(N,\theta)$ the Binomial distribution. Since the Beta distribution and the Binomial distribution form a conjugate pair, a common prior distribution on θ is $\beta(\alpha,\alpha)$ which have mean and median equal to $\frac{1}{2}$. Three 'plausible' noninformative priors were listed by Berger (1985, p.89): the Bayes-Laplace prior $\beta(1,1)$, the Jeffreys prior $\beta(\frac{1}{2},\frac{1}{2})$ and the improper Haldane prior, wrote down $\beta(0,0)$, whose density is $\pi_H(\theta)=\frac{1}{\theta(1-\theta)}$ w.r.t. Lebesgue measure on $]0,1[$. If we want $\beta(\alpha,\alpha)$ with large variance, necessarily $\alpha\to 0$. Thus, we choose $\beta(\frac{1}{n},\frac{1}{n})$. The density of $\Pi_n=\beta(\frac{1}{n},\frac{1}{n})$ w.r.t. Lebesgue measure on $]0;1[$ is $\pi_n(\theta)=\frac{1}{B(\frac{1}{n},\frac{1}{n})}$

$\theta^{\frac{1}{n}-1}(1-\theta)^{\frac{1}{n}-1}$. As mentioned, e.g. by Bernardo (1979) or Lane and Sudderth (1983), there are two possible limits for $\{\beta(\frac{1}{n},\frac{1}{n})\}_{n\in\mathbb{N}^*}$: $\frac{1}{2}(\delta_0+\delta_1)$ which is the limiting measure given by the standard probability theory and Π_H which is deduced from the posterior distributions and estimators. In fact, we are going to show that it depends on if we are on $]0,1[$ or on $[0,1]$. Choosing $]0,1[$ or $[0,1]$ doesn't matter for $\beta(\frac{1}{n},\frac{1}{n})$ but it matters for the limiting distributions. Note that the Haldane prior is a Radon measure on $]0,1[$ but not on $[0,1]$ and that $\frac{1}{2}(\delta_0+\delta_1)$ is not defined on $]0,1[$. Then, we study the convergence of posterior distributions and, as did by Lehmann and Casella (1998), we look at the behaviour of estimators.

4.4.1. *Convergence on* $]0,1[$. In this section, we study the convergences on $]0,1[$ of $\{\beta(\frac{1}{n},\frac{1}{n})\}_{n\in\mathbb{N}^*}$, of the sequence of posteriors and of the sequence of estimators.

Convergence of $\{\beta(\frac{1}{n},\frac{1}{n})\}_{n\in\mathbb{N}^*}$

On $]0,1[$, $\{\beta(\frac{1}{n},\frac{1}{n})\}_{n\in\mathbb{N}^*}$ converges q-vaguely to Π_H.

Indeed, put $a_n=B(\frac{1}{n},\frac{1}{n})$, then $a_n\pi_n(\theta)=\theta^{\frac{1}{n}-1}(1-\theta)^{\frac{1}{n}-1}$ is an increasing sequence which converges to $\pi_H(\theta)=\frac{1}{\theta(1-\theta)}$. The result follows from Theorem 2.12.

Convergence of posterior distributions

On $]0,1[$,

$$\{\Pi_n(\theta|x)\}_{n\in\mathbb{N}^*} \text{ converges } q\text{-vaguely to } \Pi_H(\theta|x)=\beta(x,N-x) \tag{4}$$

with $\beta(0,N)$, resp $\beta(N,0)$, the improper measures with pdf $\pi(\theta)=\frac{(1-\theta)^{N-1}}{\theta}$, resp $\pi(\theta)=\frac{\theta^{N-1}}{1-\theta}$.

Indeed, $\{\Pi_n\}_{n\in\mathbb{N}^*}$ converges q-vaguely to Π_H and $\theta\longmapsto f(x|\theta)=\binom{N}{x}\theta^x$ $(1-\theta)^{N-x}$ is continuous on Θ so the result follows from Lemma 3.1.

In fact, for $0<x<N$, because $\beta(x,N-x)$ is proper and $\theta\longmapsto f(x|\theta)$ is continuous and positive, from Theorem 3.3, we may replace the q-vague convergence by narrow convergence in (4).

Convergence of estimators

For all n, $\mathbb{E}_{\Pi_n}(\theta|x)=\frac{1\,+\,n\,x}{2\,+\,n\,N}\xrightarrow[n\to+\infty]{}\frac{x}{N}$. So:

- If $x = 0$, $\mathbb{E}_{\Pi_n}(\theta|x = 0) \xrightarrow[n\to+\infty]{} 0$ whereas $\mathbb{E}_{\Pi_H}(\theta|x = 0) = \frac{1}{N}$.
- If $x = N$, $\mathbb{E}_{\Pi_n}(\theta|x = N) \xrightarrow[n\to+\infty]{} 1$ whereas $\mathbb{E}_{\Pi_H}(\theta|x = N) = +\infty$.
- If $0 < x < N$, $\mathbb{E}_{\Pi_n}(\theta|x) \xrightarrow[n\to+\infty]{} \frac{x}{N} = \mathbb{E}_{\Pi_H}(\theta|x)$.

For $x = 0$ and $x = N$, $\Pi_H(\cdot|x)$ is an improper measure. In this case, $\mathbb{E}_{\Pi_H}(\theta|x) = \int \theta \mathrm{d}\Pi_H(\theta|x)$.

4.4.2. *Convergence on* $[0,1]$. In this section, we study the convergences on $[0,1]$ of $\{\beta(\frac{1}{n},\frac{1}{n})\}_{n\in\mathbb{N}^*}$, of the sequence of posteriors and of the sequence of estimators.

Convergence of $\{\beta(\frac{1}{n},\frac{1}{n})\}_{n\in\mathbb{N}^*}$

On $[0,1]$, $\{\beta(\frac{1}{n},\frac{1}{n})\}_{n\in\mathbb{N}^*}$ converges narrowly to $\mathcal{U}(\{0,1\}) = \frac{1}{2}(\delta_0+\delta_1) = \Pi_{\{0,1\}}$. Indeed, $\mathrm{med}(\Pi_n) = \frac{1}{2}$ is constant and, on $]0,1[$, $\{\beta(\frac{1}{n},\frac{1}{n})\}_{n\in\mathbb{N}^*}$ converges q-vaguely to the improper measure Π_H. So by Corollary 2.9, for all $0 < t < 1$, $F_n(t) = \Pi_n([0,t[) = \Pi_n(]0,t[) \xrightarrow[n\to+\infty]{} \frac{1}{2}$. From Proposition B.14, $\{\beta(\frac{1}{n},\frac{1}{n})\}_{n\in\mathbb{N}^*}$ converges narrowly to $\Pi_{\{0,1\}}$ which has for distribution function $F(t) = \frac{1}{2}\mathbb{1}_{0<t<1} + \mathbb{1}_{t=1}$. By Theorem 2.5, $\{\beta(\frac{1}{n},\frac{1}{n})\}_{n\in\mathbb{N}^*}$ cannot converge to an other limit such as, e.g., the Haldane measure which is not a Radon measure on $[0,1]$.

Convergence of posterior distributions

The limit of the posterior distributions can be deduced from the limit of the prior distributions only for $x = 0$ and $x = N$.

- If $x = 0$, $\{\Pi_n(\theta|x = 0)\}_{n\in\mathbb{N}^*}$ converges narrowly to $\Pi_{\{0,1\}}(\theta|x = 0) = \delta_0$.
- If $x = N$, $\{\Pi_n(\theta|x = N)\}_{n\in\mathbb{N}^*}$ converges narrowly to $\Pi_{\{0,1\}}(\theta|x = N) = \delta_1$.
- If $0 < x < N$, $\{\Pi_n(\theta|x)\}_{n\in\mathbb{N}^*}$ converges narrowly to $\beta(x, N-x)$ whereas $\Pi_{\{0,1\}}(\theta|x)$ doesn't exist.

Convergence of estimators

The limit of the estimators can be deduce from the limit of the prior distributions only for $x = 0$ and $x = N$.

- If $x = 0$, $\mathbb{E}_{\Pi_n}(\theta|x = 0) \xrightarrow[n\to+\infty]{} 0 = \mathbb{E}_{\Pi_{\{0,1\}}}(\theta|x = 0)$.
- If $x = N$, $\mathbb{E}_{\Pi_n}(\theta|x = N) \xrightarrow[n\to+\infty]{} 1 = \mathbb{E}_{\Pi_{\{0,1\}}}(\theta|x = N)$.
- If $0 < x < N$, $\mathbb{E}_{\Pi_n}(\theta|x) \xrightarrow[n\to+\infty]{} \frac{x}{N}$ whereas $\mathbb{E}_{\Pi_{\{0,1\}}}(\theta|x)$ doesn't exist.

5. Jeffreys-Lindley paradox. The use of improper priors is always delicate in hypothesis testing. Consider, for example, the standard case $X|\theta \sim \mathcal{N}(\theta, 1)$ with $\theta \sim \mathcal{N}(0, n^2)$ and the point null hypothesis $H_0 : \theta = 0$. If we use the improper prior $\pi(\theta) = 1$ on H_1, i.e., if $\pi(\theta) = \frac{1}{2}\mathbb{1}_{\theta=0} + \frac{1}{2}\mathbb{1}_{\theta\neq 0}$ w.r.t. the measure $\delta_0 + \lambda_{\mathbb{R}}$, the posterior probability of H_0 is $\Pi(\theta = 0|x) = \frac{1}{1+\sqrt{2\pi}e^{x^2/2}}$ so $\Pi(\theta = 0|x) \leqslant \left[1 + \sqrt{2\pi}\right]^{-1} \approx 0.285$ whatever the data are. An alternative is to use a sequence of proper priors $\{\Pi_n\}_{n\in\mathbb{N}}$ whose pdf are $\pi_n(\theta) = \frac{1}{2}\mathbb{1}_{\theta=0} + \frac{1}{2}\mathbb{1}_{\theta\neq 0}\frac{1}{\sqrt{2\pi}n}e^{-\frac{\theta^2}{2n^2}}$ w.r.t. $\delta_0 + \lambda_{\mathbb{R}}$. With these priors, we have $\pi_n(\theta = 0|x) = \left[1 + \sqrt{\frac{1}{1+n^2}}e^{\frac{n^2x^2}{2(1+n^2)}}\right]^{-1}$ which converges to 1. This limit differs from the "noninformative" answer $\left[1 + \sqrt{2\pi}e^{x^2/2}\right]^{-1}$ and is considered as a paradox. In the light of the concept of q-vague convergence, this result is not paradoxal since, as shown in Lemma 5.1, the priors $\{\frac{1}{2}\delta_0 + \frac{1}{2}\mathcal{N}(0, n^2)\}_{n\in\mathbb{N}}$ converges vaguely to $\rho\delta_0$, and, from Lemma 5.2, the limiting posterior distribution corresponds to the posterior of the limit of the prior distribution.

LEMMA 5.1. *Consider a partition:* $\Theta = \Theta_0 \bigcup \Theta_1$ *where* $\Theta_0 = \{\theta_0\}$. *Let* $\{\widetilde{\Pi}_n\}_{n\in\mathbb{N}}$ *be a sequence of probabilities on* Θ *which converges q-vaguely to the improper measure* $\widetilde{\Pi}$ *and such that* $\widetilde{\Pi}_n(\theta_0) = \widetilde{\Pi}(\theta_0) = 0$. *Put* $\Pi_n = \rho\delta_{\theta_0} + (1-\rho)$ $\widetilde{\Pi}_n$ *where* $0 < \rho < 1$, *then* $\{\Pi_n\}_{n\in\mathbb{N}}$ *converges vaguely to* $\rho\delta_{\theta_0}$.

PROOF. From Proposition 2.2, there exists $\{a_n\}_{n\in\mathbb{N}}$ such that $\{a_n\widetilde{\Pi}_n\}_{n\in\mathbb{N}}$ converges vaguely to $\widetilde{\Pi}$. For $g \in \mathcal{C}_K$, $\Pi_n(g) = \rho g(\theta_0) + (1-\rho)\widetilde{\Pi}_n(g) = \rho g(\theta_0)$ $+ \frac{1-\rho}{a_n}a_n\widetilde{\Pi}_n(g)$. But, $a_n\widetilde{\Pi}_n(g) \xrightarrow[n\to+\infty]{} \widetilde{\Pi}(g) < \infty$. So, $\frac{1-\rho}{a_n}a_n\widetilde{\Pi}_n(g) \xrightarrow[n\to+\infty]{}$ 0 since, from Lemma 2.6, $a_n \xrightarrow[n\to+\infty]{} +\infty$. Thus, $\Pi_n(g) \xrightarrow[n\to+\infty]{} \rho g(\theta_0)$. The result follows. □

The following proposition is a generalization of the previous example.

LEMMA 5.2. *Consider the same notations and assumptions of Lemma 5.1. Moreover, assume that* $\theta \longmapsto f(x|\theta)$ *is continuous and belongs to* $\mathcal{C}_0(\Theta)$. *Then* $\{\Pi_n(\cdot|x)\}_{n\in\mathbb{N}}$ *converges narrowly to* $\Pi(\cdot|x)$.

PROOF. It is a direct consequence of Theorem 3.6. □

In the previous Lemma, we have assumed that $\theta \longmapsto f(x|\theta) \in \mathcal{C}_0(\Theta)$. We consider now the case where the limit of the likelihood $f(x|\theta)$ when θ is outside of any compact is not 0 but $f(x|\theta_0)$. In that case, the limit of the posterior probabilities is the same as the limit of the prior probabilities, as stated in the following proposition.

PROPOSITION 5.3. *Consider the same notations and assumptions of Lemma 5.1. Moreover, assume that* $\theta \longmapsto f(x|\theta)$ *is continuous and such that for all* $\varepsilon > 0$*, there exists a compact* K *such that for all* $\theta \in K^c$ $|f(x|\theta) - f(x|\theta_0)| \leqslant \varepsilon$*. Then* $\Pi_n(\theta = \theta_0|x) \xrightarrow[n\to+\infty]{} \Pi(\theta = \theta_0)$ *and* $\Pi_n(\theta \neq \theta_0|x) \xrightarrow[n\to+\infty]{} \Pi(\theta \neq \theta_0)$*.*

PROOF. By Bayes formula: $\Pi_n(\theta = \theta_0|x) = \frac{\rho f(x|\theta_0)}{\rho f(x|\theta_0) + (1-\rho)\int_\Theta f(x|\theta)\mathrm{d}\widetilde{\Pi}_n(\theta)}$. But, for all $\varepsilon > 0$, there exists a compact K such that, for all $\theta \in K^c$, $|f(x|\theta) - f(x|\theta_0)| \leqslant \varepsilon$. So $\int_\Theta f(x|\theta)\mathrm{d}\widetilde{\Pi}_n(\theta) = \int_K f(x|\theta)\mathrm{d}\widetilde{\Pi}_n(\theta) + \int_{K^c} f(x|\theta)$ $\mathrm{d}\widetilde{\Pi}_n(\theta)$, where:

- $(f(x|\theta_0) - \varepsilon)\, \widetilde{\Pi}_n(K^c) \leqslant \int_{K^c} f(x|\theta)\mathrm{d}\widetilde{\Pi}_n(\theta) \leqslant (f(x|\theta_0) + \varepsilon)\, \widetilde{\Pi}_n(K^c)$. From Proposition 2.8, $\widetilde{\Pi}_n(K^c) \xrightarrow[n\to+\infty]{} 1$. So, $\int_{K^c} f(x|\theta)\mathrm{d}\widetilde{\Pi}_n(\theta) \xrightarrow[n\to\infty]{} f(x|\theta_0)$.
- There exists $g \in \mathcal{C}_K(\Theta)$ such that $0 \leqslant g \leqslant 1$ and $g\mathbb{1}_K = 1$. For a such g, $\int_K f(x|\theta)\mathrm{d}\widetilde{\Pi}_n(\theta) \leqslant \frac{1}{a_n} a_n \int_\Theta g(\theta) f(x|\theta)\mathrm{d}\widetilde{\Pi}_n(\theta) \xrightarrow[n\to+\infty]{} 0$ since $a_n \int_\Theta g(\theta)\ f(x|\theta)\mathrm{d}\widetilde{\Pi}_n(\theta) \xrightarrow[n\to+\infty]{} \int_\Theta g(\theta) f(x|\theta)\mathrm{d}\widetilde{\Pi}(\theta) < +\infty$ and $a_n \xrightarrow[n\to+\infty]{} +\infty$ from Lemma 2.6.

Thus, $\Pi_n(\theta = \theta_0|x) \xrightarrow[n\to\infty]{} \frac{\rho f(x|\theta_0)}{\rho f(x|\theta_0) + (1-\rho) f(x|\theta_0)} = \rho = \Pi(\theta = \theta_0)$. □

To illustrate this result in a more general case, we consider an example proposed by Dauxois et al (2006). They consider a model choice between $\mathcal{P}(m)$ the Poisson distribution, $\mathcal{B}(N, m)$ the binomial distribution and $\mathcal{NB}(N, m)$ the negative binomial distribution. These models belong to the general framework of Natural Exponentiel Families (NEFs) and are determined by their variance funtion $V(m) = am^2 + m$ where m is the mean parameter. Thus, a null value for a relates to the Poisson NEF, a negative one to the binomial NEF and a positive one to the negative binomial NEF. The prior distribution chosen on the parameter a is Π_K defined by

$$\Pi_K(a) = \begin{cases} \frac{1}{3} \text{ if } a = 0 \\ \frac{1}{3K} \text{ if } \frac{1}{a} \in \{1, ..., K\} \\ \frac{1}{3K} \text{ if } -\frac{1}{a} \in \{n_0, ..., n_0 + K - 1\} \end{cases}$$

where K is an hyperparameter. Note that $\Pi_K(a = 0) = \Pi_K(a > 0) = \Pi_K(a < 0) = \frac{1}{3}$.

Dauxois et al (2006) showed that the posterior distributions does not converge to $\delta_{\{a=0\}}$ as in the previous case but $\Pi_K(a = 0|X = x)$, $\Pi_K(a > 0|X = x)$ and $\Pi_K(a < 0|X = x)$ converge to the prior probabilities $\Pi_K(a = 0)$, $\Pi_K(a > 0)$ and $\Pi_K(a < 0)$ whatever the data are when $K \to +\infty$.

APPENDIX A: PROPERTIES OF THE QUOTIENT SPACE

We denote by $\mathcal{R}$ the space of non-zero positive Radon measures on Θ and by $\overline{\mathcal{R}}$ the quotient space of $\mathcal{R}$ w.r.t. the equivalence relation $\sim$ defined in (2). In that follows, we consider the open sets relatively to the vague topology.

Definition A.1. *The quotient topology on $\overline{\mathcal{R}}$ is the finest topology on $\overline{\mathcal{R}}$ for which*

$$\begin{array}{cccc} \phi & : & \mathcal{R} & \rightarrow & \overline{\mathcal{R}} \\ & & \Pi & \mapsto & \overline{\Pi} \end{array} \tag{5}$$

is continuous. If $\mathcal{O} = \{$open set of $\mathcal{R}\}$ then $\overline{\mathcal{O}} = \{\overline{O} = \phi(O)$ such that $O \in \mathcal{O}\}$ is the set of open sets in the quotient space.

Proof of Proposition 2.2.

- Direct part: Assume that $\overline{\Pi}_n \xrightarrow[n\to+\infty]{} \overline{\Pi}$. $\mathcal{R}$ is a metrisable space so $\mathcal{R}$ has a countable neighbourhood base. Thus, there exists a decreasing sequence of open sets $\{O_i\}_{i\in\mathbb{N}}$ in $\mathcal{R}$ such that for all $i \in \mathbb{N}$, $\Pi \in O_i$ and $\bigcap_{i\in\mathbb{N}} O_i = \{\Pi\}$. So, for all $i \in \mathbb{N}$, we have $\overline{\Pi} \in \overline{O}_i$. For any O_i, there exists N_i such that for all $n > N_i$, $\overline{\Pi}_n \in \overline{O}_i$. Without lost of generality, we can choose N_i such that $N_i > N_{i-1}$. For all n such that $N_i \leqslant n < N_{i+1}$, $\Pi_n \in \mathcal{C}(O_i)$ where $\mathcal{C}(O_i) = \{\lambda x$ with $\lambda > 0$ and $x \in O_i\}$; i.e. for all n such that $N_i \leqslant n < N_{i+1}$, there exists $a_n > 0$ such that $a_n\Pi_n \in O_i$. Moreover, since $\bigcap_{i\in\mathbb{N}} O_i = \{\Pi\}$, $a_n\Pi_n \xrightarrow[n\to+\infty]{} \Pi$.
- Converse part: Assume that $\{a_n\Pi_n\}_{n\in\mathbb{N}}$ converges to Π. Since the application ϕ defined in (5) is continuous, $\{\phi(a_n\Pi_n)\} = \{\overline{\Pi}_n\}$ converges to $\phi(\Pi) = \overline{\Pi}$.

Proposition A.2. *$\overline{\mathcal{R}}$ is a Hausdorff space.*

Proof. This proof is based on two results of Bourbaki (1971).

Step 1: $\mathcal{R}$ is a topological space and $\Gamma = \{\sigma_\alpha : \Pi \longmapsto \alpha\Pi,\ \alpha \in \mathbb{R}_+^*\}$ is a homeomorphism group of $\mathcal{R}$. We consider the equivalence relation: $\Pi \sim \Pi'$ $\Longleftrightarrow$ there exists $\alpha > 0$ such that $\Pi = \alpha\Pi'$ i.e. there exists $\sigma_\alpha \in \Gamma$ such that $\Pi = \sigma_\alpha(\Pi')$. So, from Bourbaki (1971, section I.31), $\sim$ is open.

Step 2: Let us show that $G = \{(\Pi, \alpha\Pi),\ (\Pi, \alpha\Pi) \in \mathcal{R} \times \mathcal{R}\}$ which is the graph of $\sim$ is closed. Let $\{(\Pi_n, \alpha_n\Pi_n)\}_{n\in\mathbb{N}^*}$ be a sequence in G such that $(\Pi_n, \alpha_n\Pi_n) \xrightarrow[n\to+\infty]{} (\Pi_0, \Pi_0')$. The aim is to show that $(\Pi_0, \Pi_0') \in G$, i.e. (Π_0, Π_0') takes the form $(\Pi_0, \alpha_0\Pi_0)$ where $\alpha_0\Pi_0 \neq 0$. Since $\Pi_0 \neq 0$, there exists $f_0 \in \mathcal{C}_K$ such that $\Pi_0(f_0) > 0$. Moreover, $\Pi_n(f_0) \xrightarrow[n\to+\infty]{} \Pi_0(f_0)$ so there exists $N \in \mathbb{N}^*$ such that for all $n \geqslant N$, $\Pi_n(f_0) > 0$. For all $n \geqslant N$, $\alpha_n = \frac{\alpha_n \Pi_n(f_0)}{\Pi_n(f_0)} \xrightarrow[n\to+\infty]{} \frac{\Pi'_0(f_0)}{\Pi_0(f_0)} = \alpha_0$. Thus, for all $f \in \mathcal{C}_K$, $\alpha_n \Pi_n(f) \xrightarrow[n\to+\infty]{} \alpha_0 \Pi_0(f)$ and $\alpha_n \Pi_n(f) \xrightarrow[n\to+\infty]{} \Pi'_0(f)$. Since $\mathcal{R}$ is a Hausdorff space, $\alpha_0 \Pi_0(f) = \Pi'_0(f)$. So, the graph of $\sim$, G, is closed. The result follows from Bourbaki (1971, section I.55).

□

APPENDIX B: REMINDERS ON TOPOLOGY

B.1. Some definitions and properties.

Definition B.1 (Bourbaki, 1971, p76). *A Hausdorff space is a topological space in which two distinct points have disjoint neighbourhoods.*

Lemma B.2 (Malliavin, 1982, chapter 2.2). *For a locally compact Hausdorff space that is second countable E, there exists a sequence of compact sets $\{K_n\}_{n\in\mathbb{N}}$ such that $E = \bigcup_{n\in\mathbb{N}} K_n$ and $K_n \subset \mathring{K}_{n+1}$ where $\mathring{K}_n$ is the interior of K_n.*

Lemma B.3 (Malliavin, 1982, chapter 2.2). *Let E be a locally compact Hausdorff space that is second countable. For all compact $K \subset \left(\bigcup_{n\in\mathbb{N}} \mathring{K}_n\right) = E$, there exists $N \in \mathbb{N}$ such that $K \subset \mathring{K}_N$, where $\{K_n\}_{n\in\mathbb{N}}$ is a compact sequence defined in Lemma B.2.*

Lemma B.4. *Let E be a locally compact Hausdorff space that is second countable, for all compact $K_0 \subset \left(\bigcup_{n\in\mathbb{N}^*} \mathring{K}_n\right) = E$, there exists a function $h \in \mathcal{C}_K(E)$ such that $\mathbb{1}_{K_0} \leqslant h \leqslant 1$.*

Proof. It is a consequence of Lemma B.3 and Urysohn's lemma. □

Proposition B.5 (Malliavin, 1982, chapter 2.2). *Let E be a locally compact Hausdorff space that is second countable and let μ be a positive real-valued Radon measure on E. The mesure μ satisfy the three following assertions:*

- *For all borelian set B, $\mu(B) = \sup\{\mu(K),\, K \subset B,\, K$ compact $\} = \inf\{\mu(O),\, B \subset O,\, O$ open set $\}$*

- *If $O \subset E$ is an open set and $T(O) = \{f \in \mathcal{C}_K(E, \mathbb{R})$, $\mathrm{Supp}(f) \subset O$ and $0 \leqslant f \leqslant 1\}$, then $\mu(O) = \sup_{f \in T(O)} \int f \mathrm{d}\mu$.*
- *If ν is an other positive real-valued Radon measure on E such that $\mu(f) = \nu(f)$ for all $f \in \mathcal{C}_K(E)$ then $\mu = \nu$.*

LEMMA B.6. *If the sequence $\{\mu_n\}_{n\in\mathbb{N}}$ of positive Radon measures on the locally compact space E converges vaguely to the positive Radon measure μ, then the associated total mass satisfy $\mu(E) \leqslant \liminf \mu_n(E)$.*

B.2. Reminders on convergence mode of measures. It is useful to recall the two classic kinds of convergence of measures. The following definitions and propositions are in Malliavin (1982, p91-92).

DEFINITION B.7. *Let $\{\Pi_n\}_{n\in\mathbb{N}}$ be a sequence in $\mathcal{M}^b$. We say that the sequence $\{\Pi_n\}_{n\in\mathbb{N}}$ converges narrowly to $\Pi \in \mathcal{M}^b$ if, for every function ϕ in $\mathcal{C}_b(\Theta)$, $\{\Pi_n(\phi)\}_{n\in\mathbb{N}}$ converges to $\Pi(\phi)$.*

DEFINITION B.8. *Let $\{\Pi_n\}_{n\in\mathbb{N}}$ be a sequence in $\mathcal{R}$. We say that $\{\Pi_n\}_{n\in\mathbb{N}}$ converges vaguely to $\Pi \in \mathcal{R}$ if, for every function ϕ in $\mathcal{C}_K(\Theta)$, $\{\Pi_n(\phi)\}_{n\in\mathbb{N}}$ converges to $\Pi(\phi)$.*

LEMMA B.9. *If $\{\Pi_n\}_{n\in\mathbb{N}}$ is a sequence of probability measures and Π is a proper measure, then $\{\Pi_n\}_{n\in\mathbb{N}}$ converges vaguely to Π iff for all $g \in \mathcal{C}_0(\Theta)$, $\{\Pi_n(g)\}_{n\in\mathbb{N}}$ converges to $\Pi(g)$.*

PROOF. Consider $g \in \mathcal{C}_0(\Theta)$. For all $\varepsilon > 0$, there exists a compact K_0 such that, for all $\theta \in K_0^c$, $|g(\theta)| \leqslant \varepsilon$. From Lemma B.4, there exists $g_1 \in \mathcal{C}_K(\Theta)$ such that, for all $\theta \in K_0$, $g_1(\theta) = g(\theta)$. Put $g_2 = g - g_1$. For all n, $|\Pi_n(g) - \Pi(g)| \leqslant |\Pi_n(g_1) - \Pi(g_1)| + |\Pi_n(g_2)| + |\Pi(g_2)|$. But $\{\Pi_n(g_1)\}_{n\in\mathbb{N}}$ converges to $\Pi(g_1)$ so there exists N such that for all $n > N$, $|\Pi_n(g_1) - \Pi(g_1)| < \varepsilon$. Moreover, $|\Pi_n(g_2)| \leqslant \varepsilon \Pi_n(\Theta) = \varepsilon$ and $|\Pi(g_2)| \leqslant \varepsilon \Pi(\Theta)$ with $\Pi(\Theta) < +\infty$. So, for $n > N$, $|\Pi_n(g) - \Pi(g)| \leqslant \varepsilon(2 + \Pi(\Theta))$. The result follows. □

LEMMA B.10. *Let $\{\Pi_n\}_{n\in\mathbb{N}}$ and Π be measures on $\Theta = \{\theta_i\}_{i\in I}$, $I \subset \mathbb{N}$. The sequence $\{\Pi_n\}_{n\in\mathbb{N}}$ converges vaguely to Π iff for all $\theta \in \Theta$, $\{\Pi_n(\theta)\}_{n\in\mathbb{N}}$ converges to $\Pi(\theta)$.*

LEMMA B.11. *Let $\{\Pi_n\}_{n\in\mathbb{N}}$ and Π be measures on Θ. The sequence $\{\Pi_n\}_{n\in\mathbb{N}}$ converges vaguely to Π iff, for all $\phi \in \mathcal{C}_K^+(\Theta)$, $\{\Pi_n(\phi)\}_{n\in\mathbb{N}}$ converges to $\Pi(\phi)$.*

Proposition B.12. *Let $\{\Pi_n\}_{n\in\mathbb{N}}$ be a sequence in $\mathcal{M}^b$. If $\{\Pi_n\}_{n\in\mathbb{N}}$ converges narrowly to $\Pi \in \mathcal{M}^b$ then $\{\Pi_n\}_{n\in\mathbb{N}}$ converges vaguely to Π.*

Proposition B.13. *If $\{\Pi_n\}_{n\in\mathbb{N}}$ and Π are probability measures, then vague and narrow convergences of $\{\Pi_n\}_{n\in N}$ to Π are equivalent.*

Proposition B.14 (Billingsley, 1986). *Let $\{\Pi_n\}_{n\in\mathbb{N}}$ be a sequence of probabilities and Π be a probability. If $\lim_{n\to\infty} F_n(x) = F(x)$ for each x at which F is continuous, where F_n, resp F, is the distribution function of Π_n, resp Π, then $\{\Pi_n\}_{n\in\mathbb{N}}$ converges narrowly to Π.*

Definition B.15. *A sequence of bounded measures $\{\Pi_n\}_{n\in\mathbb{N}}$ is said to be tight if, for each $\varepsilon > 0$, there exists a compact set K such that, for all n, $\Pi_n(K^c) < \varepsilon$.*

Theorem B.16 (Prohorov theorem). *If $\{\Pi_n\}_{n\in\mathbb{N}}$ is a tight sequence of probability measures, then there exists a subsequence $\{\Pi_{n_k}\}_{k\in\mathbb{N}}$ which converges narrowly to a probability measure.*

Lemma B.17 (Billingsley, 1986, p346). *If $\{\Pi_n\}_{n\in\mathbb{N}}$ is a tight sequence of probability measures, and if each subsequence that converges narrowly at all converges narrowly to the probability measure Π, then $\{\Pi_n\}_{n\in\mathbb{N}}$ converges narrowly to Π.*

Definition B.18 (Billingsley, 1968, p32). *A family $\mathcal{F}$ of random variables is called uniformly integrable if given $\varepsilon > 0$, there exists $M \in [0,\infty)$ such that $\mathbb{E}(|X|\mathbf{1}_{|X|\geq M}) \leq \varepsilon$ for all $X \in \mathcal{F}$.*

Proposition B.19 (Billingsley, 1968, p32). *If $sup(\mathbb{E}(|X_n|)^{1+\varepsilon}) < +\infty$ for some positive ε, then, $\{X_n\}_{n\in\mathbb{N}}$ is a uniformly integrable family.*

Acknowledgment. The authors are grateful to Prs. C. P. Robert, J. Rousseau and S. Dachian for helpful discussions.

REFERENCES

Barndorff-Nielsen, O. (1978). *Information and exponential families in statistical theory*. John Wiley & Sons Ltd., Chichester.

Berger, J. O. (1985). *Statistical decision theory and Bayesian analysis*, second ed. *Springer Series in Statistics*. Springer-Verlag, New York.

Berger, J. O., Bernardo, J. M. and Sun, D. (2009). The formal definition of reference priors. *Ann. Statist.* **37** 905–938.

Bernardo, J.-M. (1979). Reference posterior distributions for Bayesian inference. *J. Roy. Statist. Soc. Ser. B* **41** 113–147.

BILLINGSLEY, P. (1968). *Convergence of probability measures*. John Wiley & Sons Inc., New York.

BILLINGSLEY, P. (1986). *Probability and measure*, second ed. *Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics*. John Wiley & Sons Inc., New York.

BOURBAKI, N. (1971). *Éléments de mathématique. Topologie générale. Chapitres 1 à 4*. Hermann, Paris.

DAUXOIS, J.-Y., DRUILHET, P. and POMMERET, D. (2006). A Bayesian choice between Poisson, binomial and negative binomial models. *Test* **15** 423–432.

DIACONIS, P. and YLVISAKER, D. (1979). Conjugate priors for exponential families. *Ann. Statist.* **7** 269–281.

DRUILHET, P. and POMMERET, D. (2012). Invariant conjugate analysis for exponential families. *Bayesian Anal.* **7** 903–916.

EATON, M. L. (1989). *Group invariance applications in statistics. NSF-CBMS Regional Conference Series in Probability and Statistics, 1*. Institute of Mathematical Statistics, Hayward, CA.

HARTIGAN, J. A. (1996). Locally uniform prior distributions. *Ann. Statist.* **24** 160–173.

LANE, D. A. and SUDDERTH, W. D. (1983). Coherent and continuous inference. *Ann. Statist.* **11** 114–120.

LEHMANN, E. L. and CASELLA, G. (1998). *Theory of point estimation*, second ed. *Springer Texts in Statistics*. Springer-Verlag, New York.

MALLIAVIN, P. (1982). *Intégration et probabilités. Analyse de Fourier et analyse spectrale*. Masson, Paris.

MORRIS, C. N. (1983). Natural exponential families with quadratic variance functions: statistical theory. *Ann. Statist.* **11** 515–529.

ROBERT, C. P. (2001). *The Bayesian choice*, second ed. *Springer Texts in Statistics*. Springer-Verlag, New York.

TUYL, F., GERLACH, R. and MENGERSEN, K. (2009). Posterior predictive arguments in favor of the Bayes-Laplace prior as the consensus prior for binomial and multinomial parameters. *Bayesian Anal.* **4** 151–158.

LABORATOIRE DE MATHÉMATIQUES,
UMR CNRS 6620,
CLERMONT UNIVERSITÉ,
UNIVERSITÉ BLAISE PASCAL,
63177 AUBIERE CEDEX, FRANCE
E-MAIL: Christele.Bioche@math.univ-bpclermont.fr
E-MAIL: Pierre.DRUILHET@univ-bpclermont.fr